Aalto University
School of Business

Assessment of Systemic Risk Measures

Master's Thesis

Ville Potka

Aalto University School of Business

Economics

Spring 2017

Aalto-yliopisto
Kauppakorkeakoulu

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi

Maisterin tutkinnon tutkielman tiivistelmä

Tekijä Ville Potka

Työn nimi Systeemisen riskin mittausmenetelmien arviointi

Tutkinto Kauppatieteiden maisteri

Koulutusohjelma Taloustiede (Economics)

Työn ohjaaja(t) Juuso Välimäki

Hyväksymisvuosi 2017 **Sivumäärä** 66 **Kieli** englanti

Tiivistelmä

Systeeminen riski on noussut keskeiseksi finanssijärjestelmän sääntelyyn liittyväksi haasteeksi. Finanssialan instituutioiden toiminnan tiedetään voivan aiheuttaa riskejä taloudelle yleisesti, ja näiden riskien teoreettisesta tuntemuksesta huolimatta niiden mittaamiseen ja sääntelyyn liittyy edelleen huomattavia haasteita.

Tässä maisterin tutkinnon tutkielmassa tarkastelen systeemisen riskin mittareiden käyttökelpoisuuden arviointia. Systeemisen riskin mittaamista ja siihen liittyvää teoriaa on käsitelty kirjallisuudessa runsaasti, mutta ehdotettujen riskimittareiden arviointia on toistaiseksi toteutettu vähän. Arviointia vaikeuttavat mittausmenetelmien ajoittain heikko yhteys talousteoriaan, mittareiden monimuotoisuus sekä ekonometriset haasteet ja datan puute.

Teen katsauksen systeemistä riskiä käsittelevään kirjallisuuteen näiden haasteiden ymmärtämiseksi. Esittelen myös erään tavan mitata systeemistä riskiä SRISK-menetelmällä, ja arvioin tätä menetelmää empiirisin keinoin. Tutkielmani keskeinen havainto on, että SRISK-menetelmän mahdollisesta käyttökelpoisuudesta huolimatta sitä rajoittaa huomattava herkkyys mallinnusvalintojen suhteen ja syvällisempi ymmärrys tästä menetelmästä on tarpeen, ennen kuin sitä voidaan hyödyntää systeemisen riskin sääntelyssä.

Avainsanat systeeminen riski, makrovakausvalvonta, pankkikriisit, riskien mittausmenetelmät

Aalto University
School of Business

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of master's thesis

Author Ville Potka	
Title of thesis Assessment of Systemic Risk Measures	
Degree Master of Science (Economics and Business Administration)	
Degree programme Economics	
Thesis advisor(s) Juuso Välimäki	
Year of approval 2017	**Number of pages** 66 **Language** English

Abstract

Systemic risk in the financial system presents a daunting challenge for the regulators of the industry worldwide. The behavior of financial institutions may create risks for the functioning of the whole economy, and in spite of theoretical understanding of these risks, their measurement and regulation still faces considerable challenges.

This thesis studies the empirical assessment of systemic risk measures. A wealth of literature on the theory and measurement of systemic risk exists, yet there is little work on the assessment of the proposed measures. The assessment is challenging because of a gap between proposed measures and economic theory, diversity of these measures, as well as econometric issues and lack of data.

I review existing literature on systemic risk in order to better understand these challenges. I also present an empirical evaluation of one particular method of estimating systemic risk with the SRISK measure, and point out its shortcomings. My main finding is that while the SRISK measure may be useful for the measurement of systemic risk, its usefulness is limited by the sensitivity to underlying modeling choices, and a more thorough empirical understanding of SRISK is necessary before it can be considered as a tool for systemic risk regulation.

Keywords systemic risk, macroprudential regulation, financial crises, risk measures

Contents

List of Tables

List of Figures

1 Introduction

The measurement and management of systemic risk has attracted considerable attention in the aftermath of the global financial crisis of 2007–2009. Usually the discussion of systemic risk focuses on the possibility of a failure of a financial institution spreading to others and causing a negative effect on the real economy, but it is still a concept that lacks a unified definition (Benoit et al., 2017).

Measuring and managing systemic risk is a timely topic for regulators and the financial industry. The prevalent approach to mitigating systemic risk has been microprudential (i.e. firm-level) regulation that might not be sufficient for dealing with shocks that affect the system as a whole (Acharya et al., 2017). This presents a need for a well-defined and practically relevant way of measuring an institution's contribution to the overall systemic risk.

The primary justification for managing and regulating systemic risk comes from the negative externality it imposes on both the financial system and the real economy. The global financial crisis showed that a crisis in the financial industry can lead to a systemic shock that has considerable effects on the real economy as well, and the measures of systemic risk attempt to capture the firm-level contributions to the risk of this kind of systemic crisis. If successful, a measure of systemic risk will link an individual firm with the system as a whole, and not only assess the firm in isolation as most microprudential measures do.

In this thesis, I examine the methods developed for measuring systemic risk. I approach this topic both from a theoretical and an empirical perspective, although my focus is in the empirical validation of the proposed systemic risk measures. In particular, I am interested in capital shortfall –based measures, as they have received the most attention in the literature and their clear intuition and simple calculation makes them promising from a real-world regulatory perspective. Despite their simplicity, they can still be extended to incorporate more complex features, for example in the case of the SRISK measure by Brownlees and Engle (2017). The SRISK measure, calculated for an individual institution, attempts to capture the expected capital shortfall – in other words, required bailout funds – that a firm would incur in the event of a systemic crisis. SRISK can also be aggregated for the whole financial system in order to estimate the total costs of a systemic crisis.

I begin my thesis by reviewing the literature and theoretical framework of systemic risk in section 2, and then go on to present my own empirical methodology for examining the SRISK measure in section 3. Section 4 presents my results, and in section 5, I discuss the results in the context of the existing literature. Section 6 concludes this thesis.

2 Theoretical Framework

While lacking a unified definition, systemic risk is still described fairly uniformly in the literature. In this work, I consider systemic risk to be the risk that adverse financial conditions of one financial institution spread to others through contagion effects and consequently have negative effects on the real economy. The latter part of this definition is sometimes dropped (Benoit et al.). More specifically, I characterize the "adverse conditions" as capital shortfall which will be discussed in more detail in section 3.1. This approach is consistent with e.g. Acharya et al. (2012). Another characteristic of systemic risk is that it can be – and for regulatory purposes, should be – decomposed into contributions of individual institutions. This notion is closely related to the theory of systemic risk measurement, which I discuss in section 2.3.1. It should also be noted that while this work and the literature I survey concern systemic risk in the financial system, the concept itself is applicable to other kinds of systems as well (Chen et al., 2013).

Risks of financial institutions and of the whole system are frequently discussed in terms of categories such as market risk, liquidity risk, credit risk, etc. Systemic risk differs from this kind of source-based categorization in the way that it encompasses risks from different sources, and the focus is not so much on where the risk "comes from" but on the mechanism and possible consequences, i.e. contagion in the financial system and spillover to the real economy. The actual losses that trigger the realization of systemic risk may be caused by market price movements (market risk) or freezing of the markets (liquidity risk), for example.

An important characteristic of systemic risk is that actions by individual financial institutions create a negative externality on the real economy that is not internalized by the market. The reasons for this have been studied extensively in the literature. Benoit et al. provide a review on the topic, and I follow their exposition in the following section. But first, in order to better illustrate the framework and sources of systemic risk, I present a conceptual model described by Benoit et al.

The financial system consists of N financial institutions[1], which are indexed with i. Each of the institutions faces a risk exposure x_i. Of this exposure, the proportion α_i is considered systematic and the proportion $1 - \alpha_i$ is idiosyncratic. A systematic shock affects all the institutions at the same time and can

[1]For the purposes of this thesis, the terms financial institution and bank can be considered interchangeable, unless stated otherwise.

be understood, for example, in the context of the CAPM framework.[2] The systematic exposure of institution i is thus defined as $y_i^S = \alpha_i x_i$, and the idiosyncratic exposure as $y_i^I = (1 - \alpha_i)x_i$. For the whole system, the aggregate exposure to systematic risk is denoted $y^S = \sum_{i=1}^{N} y_i^S$. There are links between the institutions, given by the $N \times N$ matrix $\mathbf{B}$. The elements $b_{i,j}$ denote the amount of i's exposure to j, in practice in the form of derivatives contracts or interbank loans, for example.

As usual, the counterpart of risk is return. In this framework, just like risks, returns can also be attributed to systematic and idiosyncratic factors. The return on the systematic factor is denoted by $r^S + \varepsilon^S$, where r^S is a constant and ε^S is a random variable with zero mean. Analogously, the returns on the idiosyncratic factor can be denoted $r^i + \varepsilon^i$, with ε^S and ε^i being independently distributed.

A benchmark payoff (or profit) $\hat{\pi}_i$, corresponding to the payoff of an institution that is the only one in the system, can be described as a function $\hat{\pi}_i(y_i^S, y_i^I, \varepsilon^S, \varepsilon^i)$. For example, the payoff could be specified as:

$$\hat{\pi}_i = (r^S + \varepsilon^S) \times y_i^S + (r^i + \varepsilon^i) \times y_i^I \tag{1}$$

In this kind of framework, a systematic shock would be captured by ε^S, and the risk related to that factor can be called *systematic* risk.

However, systemic risk is more than only systematic risk. The institution i belongs to a system (or network) of financial institutions that can be described with $\mathbf{B}$. Systemic risk is characterized by the actual payoff π_i being different from the benchmark payoff $\hat{\pi}_i$. The links between financial institutions makes the payoff for i dependent on the exposures of other financial institutions and the idiosyncratic shocks they face. Formally, with the $N \times 1$ vectors $\boldsymbol{\varepsilon}^{\mathbf{I}}$, $\mathbf{y}^{\mathbf{S}}$, and $\mathbf{y}^{\mathbf{I}}$ denoting the idiosyncratic shocks, systematic exposures, and idiosyncratic exposures, respectively, of all the N institutions, the actual payoff π_i can be described as follows:

$$\pi_i = \pi_i(\mathbf{y}^{\mathbf{S}}, \mathbf{y}^{\mathbf{I}}, \mathbf{B}, \varepsilon^S, \boldsymbol{\varepsilon}^{\mathbf{I}}). \tag{2}$$

Sources of systemic risk are, in theory, captured by identifying the determinants of the joint distribution of π_i, i.e. which factors affect the set of realized

[2] CAPM refers to the Capital Asset Pricing Model which describes asset returns as a function of i.a. a systematic market risk factor. The model is explained briefly in section 3.1.

payoffs in the systemic context. The actual systemic risk realization, or systemic crisis, may be defined as the sum of all the π_i falling below a certain threshold, for example.

2.1 Sources of Systemic Risk

Following Benoit et al., three main categories of the mechanisms behind systemic risk can be identified: systemic risk-taking, contagion effects, and amplification effects. I describe these mechanisms in the context of the above framework and summarize the literature, following Benoit et al., on each one of them.

2.1.1 Systemic Risk-Taking

Systemic risk-taking occurs when different institutions end up choosing similar risks and taking large exposures. In the context of the framework above, having similar risks is interpreted as high α_i, and large exposures as high x_i. There have been several theoretical studies as to why this kind of systemic behavior arises.

A quite intuitive explanation why institutions might have similar risks is that they might have invested in the same assets. As summarized by Benoit et al., one reason for investing in the same (or very similar) assets is that a failing bank imposes a negative externality on other institutions through signaling effects or demand for investments. There is an incentive to minimize this externality by moving towards a situation where banks would either survive or fail together. Another example of herding behavior is brought about by regulation itself. Also in this scenario, banks have an incentive to fail together in the case that failures occur, so that government would have to organize a bailout for all of them. Should a bailout or a stimulus be applied, the expected positive effects would be maximized when a large number of the banks would be in the receiving end.

Systemic risk-taking may also occur in the context of liquidity risk. Bhattacharya and Gale (1987) show how banks may underinvest in liquid securities, increasing the possibility of a liquidity shortage in the market. From the welfare-maximizing point of view, some banks should invest in liquid assets in order to be able to provide funding to those institutions that might be in an acute need of liquidity. However, in equilibrium, all banks underinvest in the most liquid assets and rely on the other banks to provide liquidity if needed. When there is a liquidity shock in the economy, the whole system suffers if banks have too much invested in illiquid assets. Thus, all institutions have a high exposure to a systematic liquidity shock.

The realization of systemic risk may be connected to tail risks, i.e. very improbable events which can cause high losses. Tail risks are difficult to price because of their low probabilities that cannot be estimated very reliably (or at all). The literature summarized by Benoit et al. point out deficiencies in regulations that have incentivized banks, for example before the recent financial crisis, to accumulate tail risks into the shadow banking system, i.e. outside their own balance sheets. Another concern is that investors in the institutions that take excessive tail risks may not be able to recognize the risks and maintain discipline.

The leverage cycle is another well-studied phenomenon that may cause the exposures (x_i in the model above) of different institutions to be correlated. The idea is that during a boom, when asset prices are rising, agents in the economy are able to borrow more, as borrowing is restricted by the value of collateral. These constraints affect all institutions in a similar way. Related topics include the procyclicality of some risk management and regulatory measures, such as VaR[3], and the literature on bubbles.

2.1.2 Contagion Effects

A crucial feature of systemic risk is the possibility for contagion. In the theoretical framework, the contagion occurs through the bilateral links $b_{i,j}$ between financial institutions. An empirical criterion for contagion is that the payoffs of two institutions with bilateral links are positively correlated even in the absence of a systematic shock, formally: $\mathrm{Cov}(\pi_i, \pi_j | \varepsilon^S = 0) > 0$. The initial shock may thus be an idiosyncratic shock ε^i in one institution. In theory, the observation of contagion would of course require an observation of a causal relationship from one institution's payoffs to another's, but that requirement may be very difficult or impossible to fulfill in practice. Several mechanisms related to bilateral contagion and effects on the stability of the system have been studied in the literature.

[3]VaR, or value-at-risk, is some low quantile (e.g. 5%) of the distribution of payoff outcomes.

There have been several studies on how the whole network of links $\mathbf{B}$ affects systemic risk. As described by Allen and Gale (2000), the links don't only act as contagion channels but may also add robustness to the system by allowing banks to share risks. The trade-off between contagion and risk-sharing effects is determined by the completeness of the network, i.e. whether all institutions are linked to each other. If $b_{i,j}$ is positive for all institution pairs, the network is more robust than in the situation where some institutions are connected only through a third institution. Contagion effects may propagate through institutions, but risk-sharing is only possible for institutions that have a direct connection.

Another point of view to the literature of banking networks is to study how networks are determined by profit-maximizing institutions in equilibrium. Leitner (2005) describes a network where institutions are willing to organize private-sector bail-outs for failing institutions. In spite of the threat of contagion, banks may ex ante find it optimal to create linkages between each other to obtain implicit mutual insurance. In addition to these theoretical models, there are also empirical studies to the shapes of real-world banking networks. Many of them, as summarized by Benoit et al., are formed in a core-periphery fashion. Constructing sophisticated models of financial networks with real-world data forms a basis for measuring systemic risk in some studies, as summarized by Bisias et al. (2012).

The details of the infrastructure for interbank transactions may also contribute to the contagion effects. This strand of literature is concerned with issues such as net vs. gross settlement and central counterparties. Even if the underlying economics of the transactions remain same, infrastructure choices may have an effect on how large the actual linkages $b_{i,j}$ between individual institutions turn out to be. Zawadowski (2013) presents a model in which a single default can lead to a systemic crisis, arguing that having a central counterparty could mitigate the systemic risk.

2.1.3 Amplification Effects

The possibility of a relatively small shock having a large effect on the payoffs of the institutions is called amplification. If a shock affects an institution, it might have an incentive to behave in such a way that further amplifies the shock for other institutions. The larger the aggregate exposure to the shock, the bigger the effect:

$$\frac{\partial^2 \mathbb{E}(\pi_i)}{\partial \varepsilon^S \partial y^S} > 0.$$

An intuitive example of an amplification effect is a self-reinforcing liquidity crisis. Financial institutions may face constraints on their funding and collateral requirements, which are influenced by market prices. If market prices of some assets drop, institutions may be forced to liquidate positions on those assets, further contributing to the price decline. This vicious circle may lead to a large liquidity shock. This effect, combined with increasing margin requirements arising from the decreased liquidity, may force market participants to sell other kinds of assets, too, thus creating contagion between asset classes.

An extreme case of liquidity shock is called a market freeze. These are characterized by adverse selection and asymmetric information problems, which may lead to banks ceasing interbank lending altogether. Adverse selection may occur if banks cannot tell safe banks from risky banks, and banks that are inherently safe but illiquid may not access funding from the interbank market. This may also happen in a chain, where banks cease lending even though they know their counterparties but don't know the state of their counterparties' counterparties. An asymmetric information problem may occur if some of the market participants are not aware how valuable the assets used as collateral are.

There is a lot of literature on the inherent fragility of the financial system that amplifies shocks. The probability of bank runs, studied for example by Diamond and Dybvig (1983), is a classic example. Heavy reliance on short-term funding is also one factor that has been argued to contribute to the fragility of the system.

2.2 Regulatory Perspective

The presence of externalities creates a problem to be solved by the regulator. I now demonstrate this problem formally and also take a look on the real-world environment and practical issues in systemic risk regulation.

2.2.1 Regulator's Problem and Systemic Crisis

The regulator's problem can be described using a conceptual framework of the financial system. I follow Acharya et al. (2017) in this exposition, with slight deviations from their notation in order to keep variables consistent with the framework presented above.

Also here, the financial system consists of N banks (or financial institutions). The difference to the model above is that here there are two time periods $t = 0, 1$. In the economy, there are assets $k = 1, \ldots, K$ available, and each bank chooses to aqcuire total assets a^i, with investments s^i_k, as follows:

$$a^i = \sum_{k=1}^{K} s^i_k. \tag{3}$$

These investments can be financed with debt or equity. In the beginning, each bank has an endowment $\bar{w}^i_0$, of which w^i_0 is kept in the bank as equity capital and the rest, $\bar{w}^i_0 - w^i_0$, is paid out to the owner(s) as a dividend. For a bank, there may be one single owner or several owners, but they are assumed to be homogenous for a single bank in the context of this model. In addition to equity, the bank is able to raise debt funding, with a market value d^i. The budget constraint, which is equivalent to the basic accounting equation, is as follows:

$$w^i_0 + d^i = a^i. \tag{4}$$

When the system moves on to time 1, each asset k pays off a return r^i_k per dollar invested. A bit like above, a benchmark payoff can be defined, which this time stands for the total return from investments (bank revenue) in time 1, as follows:

$$\hat{p}^i = \sum_{k=1}^{K} r^i_k s^i_k. \tag{5}$$

However, the bank may face financial distress, the costs of which are denoted by ϕ^i. This leads to the actual market values of the bank's assets at time 1:

$$p^i = \hat{p}^i - \phi^i. \tag{6}$$

In turn, the distress costs themselves depend on the market value of the bank's assets and also the book value f^i of the banks debt d^i, as follows:

$$\phi^i = \Phi(\hat{p}^i, f^i). \tag{7}$$

The intuition here is that the bigger the bank and the more there is debt, the higher the eventual distress costs are. Distress costs may be incurred even if the bank doesn't default. The specification is restricted here so that $\phi \leq \hat{p}$, and thus $p \geq 0$.

Two defining characteristics of banks in this framework are that a portion of their debt is guaranteed by the government, implicitly or explicitly, and as described in section 2.1, they impose a systemic risk externality in the case of financial distress. Assuming that the government guarantee is public information, the book value of the debt is set according to

$$d^i = \gamma^i f^i + (1 - \gamma^i)\mathbb{E}(\min\{f^i, p^i\}), \tag{8}$$

where γ^i is the proportion of the bank's debt guaranteed by the government.

At time 1, the net worth w_1^i of the bank can be specified as

$$w_1^i = \hat{p}^i - \phi^i - f^i. \tag{9}$$

The owner of the bank, protected by limited liability, receives the amount[4] $[w_1^i > 0] \times w_1^i$. Now, the bank owner's problem is

$$\max_{w_0^i, d^i, \{s_k^i\}_k} c \times (\bar{w}_0^i - w_0^i - \tau^i) + \mathbb{E}(\mathrm{u}^i([w_1^i > 0] \times w_1^i)), \tag{10}$$

subject to equations (4), (5), and (7)–(9). Above, $\mathrm{u}^i([w_1^i > 0] \times w_1^i)$ is the utility for the bank owner from the time 1 income and τ^i is the bank's tax. The expression $\bar{w}_0^i - w_0^i - \tau^i$ denotes the part of the initial endowment that is paid out as a dividend to the bank's owner at time 0, and the parameter c determines the utility from taking out dividends instead of keeping the capital as equity in the bank. It can thus be understood as the opportunity cost of equity capital, which may be influenced by characteristics of the two financing methods or the bank owner itself.

Now to the regulator's problem. The aim is to maximize total welfare, given by $R = R^1 + R^2 + R^3$. The first part is the utility to all bank owners:

$$R^1 = \sum_{i=1}^{N} c \times (\bar{w}_0^i - w_0^i - \tau^i) + \mathbb{E}\left(\sum_{i=1}^{N} \mathrm{u}^i([w_1^i > 0] \times w_1^i)\right). \tag{11}$$

[4] I use the Iverson bracket notation, where $[P] = \begin{cases} 1 & \text{if } P \text{ is true;} \\ 0 & \text{otherwise.} \end{cases}$

The second part is the expected cost of the government's guarantee on bank debt, given by:

$$R^2 = \mathbb{E}\left(g \sum_{i=1}^{N} [w_1^i < 0] \gamma^i w_1^i)\right). \tag{12}$$

Here, the parameter g captures all the costs, direct and indirect, associated with this guarantee. The guarantee is paid in the case of a default, in other words when $w_1^i < 0$. The final part, R^3, is the most important one for this analysis. It is given by

$$R^3 = \mathbb{E}(e \times [W_1 < zA] \times (zA - W_1)), \tag{13}$$

where $A = \sum_{i=1}^{N} a^i$ denotes the aggregate assets in the financial system and $W_1 = \sum_{i=1}^{N} w_1^i$ is the aggregate capital in the system at time 1. If the aggregate capital W_1 falls below a certain threshold proportion z of the aggregate assets A, there is a systemic crisis. The parameter e describes the change in the severity of the crisis as aggregate capital continues to fall below the threshold given by zA. When $W_1 \geq zA$, $R^3 = 0$ and there is no systemic crisis, and a bank default in a well-capitalized system may not trigger a systemic crisis while a similar default in an undercapitalized system might. The negative externality arises from the fact that an undercapitalized system will not be able to supply credit to lending customers for their everyday business, leading to problems in the real economy. The value of the threshold parameter is essentially arbitrary, but it is needed to identify a systemic crisis. In a real-world setting, the threshold would be chosen by the regulator to best characterize a systemic crisis, possibly based on information from previous crises. It might be worthwhile to note here also that this model does not comment on the actual source (or cause) of the systemic crisis, i.e. why the capital has fallen, but it is a concern for the theoretical literature summarized in section 2.1.

The regulator's problem is to devise a tax system that maximizes total welfare, formally

$$\max_{\tau^i} R^1 + R^2 + R^3. \tag{14}$$

This means that the regulator (or government) has to be able to align incentives ex ante, so redistributions at time 1 are not possible. Actually, applying redistributions to solve a systemic crisis when it has occurred is not preferable ex ante, because it creates a moral hazard problem. Acharya et al. (2017) present a detailed analysis of optimal taxation, and conclude that the tax for an institution should be determined by two components: first, the expected loss on the institution's debt guaranteed by the government, and second, the expected costs of a systemic crisis multiplied by the institution's relative contribution to the undercapitalization. The second part is the systemic risk component, the measurement (or estimation) of which is a central topic of this thesis. In a separate contribution, Korinek (2011) shows that specifying a capital requirement is equivalent to a tax.

2.2.2 Contemporary Environment

I take a quick look on contemporary methods of regulating systemic risk, following the survey by Benoit et al.

There are several measures that could be understood in the context of the sources of systemic risk presented in section 2.1. Instead of systemic risk per se, they attempt to tackle the systemic risk-taking that may act as precursor to systemic risk, in addition to other risks. These measures[5] include capital ratios, stricter definition of capital, liquidity requirements, and countercyclical capital buffers, for example. In the light of the framework in section 2.2.1, their effect can be understood as decreasing the probability of $W_1 < zA$. However, without aligning incentives based on individual institutions' systemic risk contributions, they are not targeting the actual systemic risk externality as it is understood here. Another form of regulation are regularly conducted and published stress tests, which are more specifically attempting to identify systemically important financial institutions (SIFIs). The current regulatory framework for identifying systemically important banks is discussed in section 2.3.3.

When a crisis occurs, it will be optimal ex post if the regulator provides relief to the financial system. This is what was seen in the recent financial crisis as well. However, it is not optimal ex ante and it is a problem for the regulators to convince the financial system that there will be no bailouts in the future. Another important issue in the current regulatory environment is the international nature of crises which would require stronger international cooperation to prevent a "race to the bottom" in regulation. An example of this cooperation is the Financial Stability Board (FSB), established by the G20 in 2009 to foster cooperation and identify systemic weaknesses and SIFIs.

[5]For lack of a better word, I use the word *measure* here to refer to a regulatory policy or instrument. It should not be confused with *measure* referring to e.g. a systemic risk measure, or a unit of measurement. It should be clear from the context which interpretation of the word is intended.

2.2.3 Practical Considerations

While the theory behind systemic risk is fairly simple, there are some issues that might affect the feasibility of regulating it in the real world. One of them is the Lucas critique (Lucas, 1976): changes in policy will affect the structure of econometric models. As summarized by Bisias et al., if regulators were to supervise and measure a certain metric and base decisions on that, that would certainly affect the incentives of the financial institutions that both have an effect on the measure and are also affected by it. One particular example of this kind of effect is the shadow banking system, into which financial institutions may shift their risks in order to avoid capital requirements based on their balance sheet characteristics. Even though the banks may evade regulation this way, the risks still remain.

It can be argued that systemic risk measurement and regulation would have such benefits, especially in the presence of market imperfections, that the general-equilibrium effects of changed institutional and individual expectations would not overweigh them. This makes a case for regulating systemic risk even while acknowledging that the Lucas critique might have a bearing on this particular topic. However, any regulatory framework must take into account any possible changes in the institutions' behavior following the introduction of the regulation, such as the possibility of shifting risks into the shadow banking system.

As noted by Benoit et al., many of the tools proposed for systemic risk regulation are quite new. They can also be quite heavy-handed from the perspective of the banks if they target specific asset classes, and they may have a reducing effect on lending. This, in addition to the Lucas critique, presents a need to study and evaluate any regulation carefully before putting it into practice.

2.3 Measurement of Systemic Risk

The regulation of systemic risk by e.g. capital requirements depends on specifying a measure to capture the systemic risk contribution of an institution. I now turn to discuss how, and if, these contributions could be measured from a theoretical point of view.

2.3.1 Theoretical Foundations of Systemic Risk Measures

The systemic risk literature puts significant focus on studying and developing measures that would give regulators the possibility to capture and mitigate the systemic risk externalities of financial institutions. Keeping this in mind, any measure of systemic risk must be able to satisfy certain conditions, so that applying it as a tax or capital requirement would really fix the issue with minimal external effects of its own.

Chen et al. present an axiomatic approach to the measurement of systemic risk in order to determine the conditions for a theoretically justifiable systemic risk measure. They present a one-period framework with a set of firms $\mathcal{F}$ and a set of future scenarios $\mathbf{\Omega}$. An economy $\mathbf{X}$ is defined as a matrix of all the firms and scenarios, namely $\mathbf{X} \in \mathbb{R}^{|\mathcal{F}| \times |\mathbf{\Omega}|}$, where each element $\mathbf{X}_{i,\omega}$ denotes the loss of firm i in scenario ω. The vector of outcomes for each firm in a single scenario is $\mathbf{X}_\omega \in \mathbb{R}^{|\mathcal{F}|}$. This is referred to as a cross-sectional profile of losses. As an additional help to the exposition, unit loss vectors $\mathbf{1}_{\mathbf{\Omega}}$ and $\mathbf{1}_{\mathcal{F}}$ are defined as unit losses for one firm in all scenarios and for all firms in one scenarios, respectively. Finally, a systemic risk measure ρ is a summary statistic that quantifies the level of systemic risk in an economy X as a real number $\rho(X)$. Systemic risk measures can be compared as follows: if $\rho(X) > \rho(Y)$, the economy Y is less risky and thus preferred to X.

Chen et al. define a single-firm risk measure ρ, for all loss vectors x, y of a single firm, as satisfying the following conditions:

1. *Monotonicity*: If $x \geq y$, then $\rho(x) \geq \rho(y)$. If losses in one scenario are higher, then should the single-firm risk measure be, too.
2. *Positive homogeneity*: $\rho(x)$ increases in proportion to losses.
3. *Convexity*: Diversification reduces risk, in other words the risk of a firm diversified between outcomes x, y is less than or equal to the weighted risks of firms corresponding to x and y.
4. *Normalization*: $\rho(\mathbf{1}_{\mathbf{\Omega}}) = 1$. This criterion fixes the scaling of the risk measure.

A systemic risk measure is defined analogously. However, it introduces two additional criteria. The convexity criterion is extended from the one above to incorporate not only outcomes, but also risks. With risk aversion, combined economies yield lower uncertainty and thus increase welfare directly. The other new criterion is preference consistency that allows for comparing two cross-sectional profiles with each other. This is the main takeaway from the systemic risk perspective, as a measure satisfying this criterion will be able to order whole cross-sections according to preference relations, capturing the interaction of many institutions.

This framework allows the theoretical assessment of systemic risk measures. If a proposed measure fails to meet these criteria, its feasibility as a measure of systemic risk is questionable. For example, if an economy X with much higher expected losses than Y across the cross-section would receive a systemic risk estimate $\rho(X) < \rho(Y)$, it would not be a very useful measure if losses are what the regulator is trying to anticipate. Or, if with some measure there would be no way of determining which one of two different economies is preferable from a systemic risk point of view, the preference consistency criterion would be violated and there would be little reason to call the measure a systemic risk measure at all.

With the help of this framework, Chen et al. demonstrate that a systemic risk measure can be decomposed to the contributions of different institutions, and show that this risk attribution can account for the externalities caused by the institutions and thus serve as a basis for the tax like the one described in section 2.2.1. They also demonstrate that the risk measure can be characterized by two independent components: an aggregation function, that determines the cross-sectional preferences, and a base risk measure, that determines the preference over outcomes in different scenarios like the single-firm measure presented above.

The approach in the SRISK measure, studied in more detail in section 3.2, is closely related the systemic expected shortfall (SES) measure described briefly in the beginning of that section. The systemic expected shortfall is studied by Chen et al. as a special case of the measures included in their framework, which gives also SRISK theoretical justification.

2.3.2 Types of Measures

I now present a summary of proposed systemic risk measures. The measures can be classified in various ways, and I present some classifications found in the literature.

Benoit et al. divide systemic risk measures into structural and global measures. Following the categorization presented in section 2.1, structural measures are those that target a specific source of systemic risk, i.e. systemic risk-taking, contagion, or amplification. Systemic risk-taking has been estimated with measures like correlation in banks' portfolios and asset commonality. The latter is based on the intuition that the more banks' portfolios overlap, the more there is fragility in the system. Other ways to measure systemic risk-taking include macroeconomic models describing the economy as either in a normal or a "systemic risk state" and calculating the associated probabilities, and dynamic models studying the development of and relationship between VaR figures of GDP and financial portfolios.

The contagion effects have been measured mostly with network-theoretic methods. The measures model the linkages between institutions and study the transmission of insolvency across the network. The measures used to determine an institution's contribution to systemic risk may be e.g. distance– or centrality-based figures. Amplification effects have been quantified using structural models on e.g. leverage and liquidity, and systemic risk is measured as the sensitivity to a shock.

Global measures, on the other hand, do not target a particular source of systemic risk but attempt to capture the risk as a whole, using mostly data from the financial markets. The idea is that a lot of information can be inferred from market pricing of various securities, such as equity and derivatives. As noted by Benoit et al., global measures of systemic risk have gained a central position in the systemic risk literature, and the SRISK measure I study can be included in this category.

Other kinds of classifications for systemic risk measures are presented by Bisias et al. The measures can be classified using a temporal dimension, where a systemic risk measure can be a forecasting measure, a contemporaneous measure, or even an ex-post measure, or a combination of these. The forecasting measures are probably the most intuitive and it could even be argued that they are the most useful kind of measures for the regulator. Measures that attempt to achieve a high rate of forecasting power are typically focused on some individual metric of the economy, such as consumer credit risk or hedge-fund returns. There are also network-based measures that aim at providing an early warning of accumulating imbalances in the system. Another kind of forecasting measures include those based on stress tests, simulation, or some measure of fragility in the financial system. Many of the measures serving as basis for stress tests target a particular asset class, instead of institutions, and they attempt to capture the biggest sensitivities in the system for exogenous changes in systemic variables. The SRISK measure is an example of a measure that can be applied in conducting stress tests and identifying fragilities (Acharya et al., 2014). Identifying fragilities is also a key point in using contemporaneous measures, which attempt to provide useful and timely information for managing a crisis that has already begun. Many of the measures suitable for stress testing and short-term forecasting can also be applied for crisis monitoring. The importance of ex-post measures is mostly in the orderly resolution of failed institutions, where e.g. network-based measures may provide assistance to dealing with complex asset portfolios.

Still another way of categorizing the systemic risk measures, also described by Bisias et al., is looking at the methodological frameworks of these models. Measures like SES that attempt to capture the joint distribution of negative outcomes for financial institutions can be classified as probability-distribution measures. Default measures estimate the likelihoods of default – on consumer credit, for example – and link them across financial institutions. Techniques for these estimations include methods that are parametric as well as non-parametric, such as machine learning –based. Contingent claims analysis makes use of derivatives pricing models in estimating default probabilities. A class of their own are illiquidity measures, which treat illiquidity in the markets as an important source for systemic risk and try to estimate it using models that are compared with observed market data. Network-based measures model the connections between financial institutions using data on interbank exposures, and there have been several studies using this kind of methods to assess the externalities arising from bank failures. Commonality of asset returns across the financial system can also be analyzed with network-based measures, as well as methods such as principal component analysis. The last category described by Bisias et al. is macroeconomic measures, which differs from all the above measures by its focus on macroeconomic models and variables. The complexity of the real-world economy reduces the usefulness of macroeconomic models in estimating systemic risk.

2.3.3 Challenges

The measurement of systemic risk presents empirical and practical challenges, and some proposed measures have attracted criticism also from a theoretical point of view. In this section, I focus on shortfall-based measures, which have their foundation in the framework presented in section 2.2.1. I present the empirics of these measures in more detail in section 3.1.

The shortfall-based measures have been criticised for being atheoretical (Bisias et al.). The criticism refers particularly to their lack of connection to economic theory, not to the theoretical aspects of the measures themselves. In other words, they may have a theoretical foundation such as the one presented in section 2.2.1, but the connection to standard macroeconomic models is left unexplained. From a theoretical point of view, this is a deficiency in the model as its observations can be only interpreted in a limited context. They are also indifferent in regard to the nature of the initial shock behind the crisis, which might have a significant impact on how the crisis unfolds.

However, the lack of connection to economic theory does not necessarily mean that using shortfall-based measures is futile. As Bisias et al. note, the estimates of correlated losses produced by these models may serve as inputs for more structured systemic risk models. Also, not being "fixed" to a certain theoretical model of the economy the shortfall-based models may offer versatility and provide useful information regardless of any theoretical assumptions of the macroeconomy and the financial system.

In particular, the systemic expected shortfall (SES) has been studied in more detail by Chen et al. in the context of their axiomatic framework of systemic risk measures. As described in section 2.3.1, a systemic risk measure can be decomposed into an aggregation function and a base risk measure. In the case of SES, the base risk measure is the expected shortfall and the aggregation function can be interpreted as simply the sum of profits and losses across the cross-section of firms. SES can thus be criticized for taking an approach that treats the system like an investment portfolio and implicitly nets the profits and losses of firms with each other. This kind of an approach presents two issues. First, it might not be possible or desirable in the real world for the regulator to compensate an institution's losses with another institution's profits, even though it would be optimal from a portfolio management perspective. Second, the approach doesn't allow for a detailed modeling of preferences across cross-sections. For example, the regulator might have a view on whether it would be preferable to have a single large loss in one institution or several smaller losses in a higher number of institutions. The portfolio approach is unable to differentiate between these cases if the total magnitudes of losses are equal. These empirical limitations represent the tradeoff between theoretically simplistic models that are relatively easy to apply in practice, and complex and more structured models that may suffer from lack of transparency and real-world applicability.

Even though the new approaches may have limitations, the existing approaches for measuring and mitigating systemic risk have received criticism as well. The Basel Committee on Banking Supervision (BCBS) has a framework for analyzing the systemic importance of financial institutions (BCBS, 2013). It is a simple scoring methodology that aims to measure an institution's systemic importance by aggregating information on five categories, which determine the scores that are then scaled to reflect the relative importance of that institution in the system. Benoit et al. present a criticism of this approach. The comparison of individual institutions to the whole system presents an issue if there is a global shock that affects all institutions at the same time. The indicator may give similar values as before, even though the risk levels have changed. In the cross-section, the indicator may be useful but on the time-series dimension it might not be that reliable. Another deficiency of the BCBS approach is that the most volatile categories may dominate the changes in the risk scores. Even though the explicit goal is to give each of the five categories an equal weight, in reality those variables that change the most may determine the changes in the scores. The remedy of BCBS is to apply a cap on some of the categories, but Benoit et al. criticize it being a crude way of adjusting the scores, and propose an alternative method of scaling based on standardizing the variables based on their cross-sectional volatility.

Another common way of attempting to measure systemic risk is using stress tests that simulate macroeconomic shocks in order to calculate an institution's capital requirement. These tests make extensive use of confidential supervisory data provided by the banks. As noted in section 2.3.2, some systemic risk measures may be used as alternatives to the regulatory stress test, which allows the comparison between the approaches proposed in the literature and those used by regulators. Acharya et al. (2014) compare the results from regulatory stress tests with an alternative test based on SRISK, and find that especially in Europe, the capital required by SRISK – a market-based measure – is higher than that required by the regulatory tests, and even the ranking of institutions is different between these two tests. They also find that the ratios of risk-weighted assets to total assets in a 2011 European stress test have no link with subsequent realized bank risks or a comparable measure calculated with SRISK. They attribute the differences in the results to stress scenarios that were politically chosen to be weak and the static risk-weighting of assets. They also argue that banks have an incentive to overconcentrate their assets into those classes that have low risk weights, leading to a lack of diversification, underestimation of risk weights, and excess leverage.

In addition to theoretical challenges, there are also empirical and practical challenges in implementing the proposed systemic risk measures. Perhaps the most obvious practical problem in estimating these measures is the lack of data. While some data can be collected from financial markets or public accounting figures of financial institutions, many proposed measures require data that might be difficult to gather. A case in point are the regulatory stress tests, which rely on highly confidential data of interbank exposures. While it might not be a problem for the regulator to calculate systemic risk capital requirements from confidential data and with complex models, it is quite probable that the more the approach resembles a black box for outside observers the more it would receive opposition inside and outside the financial industry. The network-based systemic risk measures mentioned above are another example of potential methods of estimating systemic risk that would require a lot of confidential data on interbank exposures to work properly.

The complexity of the calculation of a systemic risk measure is another important consideration. While relatively simple models that only use public data may be easy to implement in the real world, they might not be able to capture the full extent of systemic risk. However, the complexity of the calculation framework may present practical obstacles if significant resources are required to come up with the figures. Any errors in calculation may also be harder to spot in complex models, and extensive use of accounting figures in the model may be hampered by differing interpretations of accounting standards across institutions. Also, if the regulator needed to be able to monitor systemic risk development for example on a daily frequency, with detailed data on interbank transactions and bank balance sheets, it would need an infrastructure for that purpose. If the cost of such an infrastructure was high compared to the benefits the model would bring in comparison to simpler systemic risk models, it might be difficult to gain acceptance for that kind of a measure.

A simple measure based on market data may not be optimal from the data requirement point of view either. Even if market data –based models were found to be effective in measuring systemic risk, they would run into problems when the institutions are privately owned and no data is available on their equity prices, for example. A significant percentage of bank assets are in unlisted institutions, as reported by e.g. Alesina and Giavazzi (2010), in whose sample the the proportion of bank assets in listed banks was 53% in Europe and 47% in the U.S. in year 2006. There is no simple remedy into this problem if only e.g. equity-based measures are used. Some of the larger privately-owned banks might have issued other traded securities, such as bonds that could provide some information on the market perception of the risks associated with these institutions, but that would require shifting from purely equity-based models to something else. Also, there are many small banks in the system that most probably haven't issued any securities on the markets. A significant systemic crisis can emerge from a number of small institutions, as evidenced by the US Savings and Loan crisis of the late 1980s, when over 1000 (mostly small) banks failed. For a more detailed account of the crisis, see e.g. Curry and Shibut (2000). These limitations are critical if the market data –based measures are to be used as the sole way of measuring systemic risk, but these measures may still be able to provide useful information if used in connection with or as an input into other measures.

In addition to the known problems and all idiosyncratic problems associated with individual research methods, a common challenge to implementing the proposed systemic risk measures in practice is the lack of their empirical evaluation. As Benoit et al. note, many of the tools are new and in order to become viable regulatory instruments, their evaluation is an important topic that deserves more attention. This thesis attempts to add especially to that particular strand of systemic risk literature. Naturally, new evaluations of the measures may also present novel challenges and questions.

Hansen (2012) presents an overview on the challenges related to systemic risk measurement. He argues that even though model misspecification is a valid concern, it is not a reason not to try and explore various proposed models and provide insights for policy advice that way. Hansen stresses the importance of sensitivity analysis when trying different models and expresses confidence that even though there may be significant challenges at the moment, studying the proposed measures will always increase the knowledge on systemic risk issues and thus bring the work forward.

2.4 Evaluation of Systemic Risk Measures

As noted by e.g. Benoit et al. and Hansen, the evaluation of systemic risk measures is an important research topic going forward. It is not only needed for regulatory purposes, but also to increase the understanding of systemic risk in general.

The rationale for the evaluation of systemic risk measures is discussed by Hansen. As the main justification for systemic risk measurement comes from the needs of regulation, the regulator's point of view is important also here. He argues that if there wasn't a commonly accepted framework for measuring systemic risk, the regulator would rely on ad hoc measures and being able to recognize systemic risk in the system when it is there. In addition to possible methodological inconsistency, too much discretion on behalf of the regulator would present problems also to the systemic risk mitigation process itself. Hansen presents two reasons for this. First, discretionary policies are vulnerable to political pressure. Second, even though discretion might allow quicker reaction in unprecedented circumstances, the criticism and development of policy would be quite challenging. In order to improve systemic risk models and methods, also other parties than the regulator should be able to participate in the systemic risk measurement discussion.

Hansen refers to the study by Bisias et al. that presents a multitude of different systemic risk measures, summarized also here in section 2.3.2. Even though the wealth of different measures presents a wide range of possible paths for improving the understanding of systemic risk, Hansen also sees it as disconcerting. As many of the proposed measures derive from very different backgrounds, any kind of systemic risk measurement framework that relies on having these several different ways to quantify systemic risk could easily suffer from lack of coherence. It is thus important to evaluate which measures work and reject those that don't.

A specific challenge that Hansen considers crucial is the differentiation of systemic and systematic risk. This distinction has been discussed also in the beginning of section 2. While systematic risk is a well-understood and well-defined concept, the nature of systemic risk is not understood that well, and distinguishing these two should be an important focus in further studies.

There seems to be no common framework for evaluating systemic risk measures. However, two main approaches can be identified: theoretical and empirical evaluations. The latter seems to be much more common in the literature, but there are also evaluations of systemic risk measures that focus on their theoretical methodology. The paper by Chen et al. is a good example of this kind of studies.

Benoit et al. summarize different approaches to empirically assessing systemic risk measures. What many of them have in common is that a systemic risk measure is studied in the context of an identified realization of systemic risk, the definition of which is different across studies. For example, the paper by Brownlees and Engle presents the SRISK measure and evaluates it by studying if institutions with a high SRISK were more likely to receive bailout funds in the financial crisis of 2007–2009. As alternatives to bailouts, the realization of systemic risk has been identified as bank insolvencies, accounting losses and negative stock returns. Another approach, also illustrated in this thesis, is comparing the rankings of several different systemic risk measures with each other and with the bucket-level rankings provided by the FSB in their annual list (see e.g. FSB, 2016). The intuition is that the regulators, who have access to confidential data, can identify systemic risk based on that data and rank the institutions accordingly. However, even if the regulator might have superior knowledge in comparison to e.g. markets and academia, their estimations may also suffer from shortcomings like the ones presented in section 2.3.3, and they should not be considered as an absolute truth.

3 Empirical Methodology

In the empirical section of this thesis, I present an empirical evaluation of the SRISK measure, comparing it with a purely systematic risk measure (CAPM beta) as well as with the list of systematically important institutions compiled by the Financial Stability Board (2016). The purpose of this approach is to determine if the SRISK appears to provide information about systemically important institutions that can't be provided by a purely market-based systematic risk measure. In addition to market data, SRISK takes into account balance sheet characteristics of financial institutions.

I have chosen to study specifically the SRISK measure for three main reasons. First, as noted by Benoit et al., SRISK is among the most studied and discussed measures in the systemic risk literature. Any further understanding on it would contribute to a growing body of work that may have significant implications for regulatory policy in the future. Even though studying a lesser-known measure might have a relatively bigger significance for the literature on that particular measure, I argue that there are justified reasons for SRISK's popularity. This brings us to the second reason why I chose SRISK for my topic. SRISK is intuitive and easy to use. SRISK fits well to the theoretical framework described in section 2.2.1, as I present below in section 3.2. It also overcomes many of the empirical challenges presented in section 2.3.3, as it relies completely on public data, and it can be understood in the context of the theoretical axioms of Chen et al. (section 2.3.1). There is also a lot of readily available data on SRISK on the website of the Volatility Institute of NYU Stern School of Business (V-Lab, 2017).

However, there is still research to be done on evaluating SRISK, which is the third reason for my interest. There are conflicting accounts on the usefulness of the SRISK measure in identifying systemic risk, and I will review the earlier assessments briefly in section 3.2.1. There are also empirical limitations and conceptual issues that may decrease the usefulness of SRISK, and I discuss those in that section as well.

3.1 Conceptual Background

In order to understand the SRISK measure, I briefly explain the shortfall-based methodology behind SRISK, as well as some other concepts needed to understand the empirical framework. In this exposition, I follow Brownlees and Engle. The notation corresponds to that in earlier sections unless stated otherwise.

The intuition behind the shortfall-based measures is described in section 2.2.1, where a capital shortfall in the economy causes a systemic crisis. The financial distress of a financial institution is also measured as capital shortfall, and for a particular firm it corresponds to the difference between the equity capital the firm needs to hold and what it actually holds. The need for a certain fraction of the firm's assets to be equity may come from e.g. regulation or prudential management practices. This prudential capital fraction is denoted k. The capital shortfall of firm i on day t is defined as

$$\begin{aligned} \mathrm{CS}_{it} &= ka_{it} - w_{it} \\ &= k(d_{it} + w_{it}) - w_{it}, \end{aligned} \tag{15}$$

where w_{it} is the market value of the firm's equity, d_{it} is the book value of debt, and $a_{it} = d_{it} + w_{it}$ is the quasi value of assets – so called as it incorporates both book and market values. If the capital shortfall is positive, the firm is in distress, and if it's negative, the firm functions properly. In the context of systemic risk measurement, the interest is in measuring the capital shortfall of an institution conditional on a systemic crisis.

The systemic crisis was defined as aggregate capital W falling below a certain threshold fraction of assets z in section 2.2.1. In this framework, a systemic crisis is characterized by an equity market decline, denoted by C. The decline happens over a time horizon h. The market return over the period from $t+1$ to $t+h$ is denoted $R_{mt+1:t+h}$, and the systemic crisis is thus the event $\{R_{mt+1:t+h} < C\}$. The derivation of the SRISK measure follows intuitively from these, and is presented in section 3.2.2.

Brownlees and Engle introduce the concept of long-run marginal expected shortfall (LRMES), which is defined as

$$\mathrm{LRMES}_{it} = -\mathbb{E}_t(R_{it+1:t+h} | R_{mt+1:t+h} < C). \tag{16}$$

It is simply the expected negative return on a firm's equity conditional on a systemic event, and is an essential component of SRISK. However, its estimation involves significant complexities as it is a forward-looking measure. Estimating future returns in the financial markets requires models of the asset price development and assumptions about the distribution of future returns. The empirical analysis in this thesis shows that the choice of model may have a significant impact on the SRISK estimation results. I discuss different approaches to estimating LRMES in section 3.2.4.

As demonstrated in the framework presented in section 2, there is a connection between systematic risk and systemic risk. Systematic risk in financial markets is often estimated with the Capital Asset Pricing Model (CAPM), and specifically its parameter β. This is also true for SRISK. The Capital Asset Pricing Model states that the return r_a of an asset a can be represented as

$$r_a = r_f + \beta(r_m - r_f), \tag{17}$$

where r_f is the risk-free rate and r_m is the (expected) market return. The parameter β can be estimated with a simple regression of the asset return above the risk-free rate on the market return above the risk-free rate, and it measures the sensitivity of the asset returns to market return. The β can be referred to as a measure of market risk or systematic risk, or simply the CAPM beta. For a more thorough description, see e.g. Howells and Bain, 2008 (pp. 192 – 195).

An important distinction to be made when studying systemic risk measures over time is between cross-section and time-series dimensions. The time-series dimension describes the evolution of a measure over time, and in the case of SRISK, can be presented for a single institution or for the whole group of institutions as an aggregate measure. The cross-sectional dimension, on the other hand, describes the value of the measure across different institutions at a certain point in time. Both dimensions are important in the empirical studies.

3.2 SRISK

The SRISK measure was initially introduced in the papers by Acharya et al. (2012) and Brownlees and Engle (2017), and it provides some improvements to earlier measures of systemic risk.

A closely related measure of systemic risk is the systemic expected shortfall (SES, Acharya et al., 2017).[6] SES is also defined with the help of the framework presented in section 2.2.1, but it differs a bit from SRISK. As explained by Brownlees and Engle, estimating SES requires an observation of systemic risk

[6] The precise definitions of SES and MES are not in the scope of this paper, nor important for the understanding of the differences between SES and SRISK, so they are left out. The article by Acharya et al. (2017) provides the definitions for those who are interested in a more detailed explanation. MES is defined analogously to LRMES, but for a time period of one day and usually with a value-at-risk interpretation of the threshold.

realization. Thus, it can't be used for ex-ante measurement, and, as shown in their paper, SRISK can predict systemic crises significantly better than SES. Another weakness of SES is that it requires assuming the structural framework described in section 2.2.1 of this paper. A benefit of SRISK is that while that framework presents an intuitive way of understanding the context and issues related to systemic risk, SRISK generalizes the notion of systemic risk so that all the assumptions of the structure of financial institutions and the system presented above are not necessary.

SRISK also incorporates more information than SES. SES is a purely market-based measure, but SRISK takes also balance sheet information into account. This allows capturing the effects the size and the leverage of the institution have on its systemic risk contribution, while still relying on public data only.

The SRISK also presents an alternative to the risk weights used in the Basel capital requirements, which Acharya et al. (2012) note as widely criticized. The Basel risk weights attempt to capture the risk in the assets an institution holds, and thus determine the capital requirement for the institution. According to Acharya et al., the LRMES component of SRISK incorporates the risks of underlying assets, and the SRISK as a whole complements the risk weight approach by taking into account the risk contribution of the institution itself.

3.2.1 Earlier Assessments and Limitations

The shortfall-based measures and also SRISK in particular have been discussed in numerous studies. In this section, I present a brief overview on earlier evaluations of the SRISK measure. I focus on the benefits and shortcomings found in SRISK, without trying to present a complete review of the empirical literature.

Obviously, the benefits of SRISK are argued for in the paper by Brownlees and Engle where SRISK is presented. They conduct two empirical tests of SRISK and find that its improvements to existing measures are not limited to theoretical aspects.

First, Brownlees and Engle study whether SRISK can be used as a predictor of capital injections by the Federal Reserve (Fed). They use a Tobit regression that determines the capital injection CI_i^* a firm i receives as follows:

$$\log \mathrm{CI}_i^* = \alpha_0 + \alpha \log(1 + (\mathrm{SRISK}_i)_+) + \boldsymbol{\gamma}'\mathbf{x}_i + \epsilon_i, \tag{18}$$

where x_+ denotes $\max\{x, 0\}$, $\mathbf{x}_i$ is a vector of control explanatory variables and ϵ_i is a Gaussian error term. The observed capital injection is thus

$$\log \mathrm{CI}_i = \begin{cases} \log \mathrm{CI}_i^* & \text{if } \log \mathrm{CI}_i^* > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{19}$$

This model is then estimated with a maximum likelihood method that finds the parameters that maximize the likelihood of making the observations, using data after March 2008. The authors find that when SRISK, SES, and ES (expected (capital) shortfall, proposed by Lehar, 2005) indicators of systemic risk are separately included in the model, SRISK improves the pseudo R^2 of the regression above the baseline the most, even though the margin is not very wide.[7] When they are included simultaneously, SRISK has the strongest statistical significance of the three. That speaks in favor of SRISK versus the other measures, but still doesn't quite prove the usefulness of SRISK. However, their conclusion from these results is that SRISK improves the prediction of Fed capital injections during the latest financial crisis.

Their second assessment of the SRISK measure uses an OLS regression model where macroeconomic variables – namely, industrial production and unemployment rate – are separately regressed on lagged values of changes in aggregate SRISK of the financial system. They find that after controlling for many economic (mostly financial) variables, changes in the aggregate SRISK predict changes in industrial production and the unemployment rate. These empirical assessments allow Brownlees and Engle to conclude that SRISK is useful in predicting systemic crises and specifically the bank-level contributions, as well as negative spillovers to the real economy.

Another assessment of the SRISK measure, conducted by Wu and Zhao (2014), studies the predictive power of the aggregate SRISK on bank failures in the United States. They study various alternative measures, such as CoVaR (Adrian and Brunnermeier, 2016), DIP (Huang et al., 2012), and a set of financial market variables, such as trailing bank equity returns, the LIBOR-OIS spread, and the TED spread which all measure the perceived risk in banks. They find support for all the measures predicting the number of bank failures. Also, a study by Brownlees et al. (2015) on data for 60 years, beginning in

[7] Pseudo R^2 is a measure of goodness-of-fit and can be understood as the logistic equivalent of the coefficient of determination (R^2) in OLS regression.

the 19th century, found SRISK useful in predicting conditions likely to cause financial crises.

An important note here is that while the aggregate SRISK may be successful in predicting economic downturns and even systemic crises and thus give support to the intuition behind it, it is not a sufficient result from the regulatory perspective. The interest of the regulator is (at least here) in measuring the systemic risk contributions of individual institutions, which the aggregate SRISK measure can't do by definition.

Idier et al. (2014) study the usefulness of the MES measure in predicting systemic risk by using the latest financial crisis as a natural experiment. They find little evidence in support of MES, and conclude that simple ratios calculated from balance sheet data fare better. They also study SRISK and find that it is not more useful than MES, which is interesting as SRISK is specifically supposed to incorporate the balance sheet metrics into the estimation. However, Idier et al. only measure realized systemic risk as realized equity losses of 19 bank holding companies, which may be a somewhat limited sample and context.

While the assessments presented here give cautious support for the usefulness of SRISK in quantifying systemic risk, there are still limitations that may hamper its use in practice. One of them is that SRISK relies on equity market data, which either calls for adjustments to account for e.g. cooperative banks or prevents its use altogether. This issue must be solved before it can be used as a basis for regulation. Also, SRISK may be criticized for not incorporating various sources of systemic risk widely enough. As presented in section 3.2.2, SRISK incorporates important characteristics that at least theoretically predict systemic risk, such as the size of the institution and its sensitivity to systemic shocks, but it implicitly relies on either market pricing in or the balance sheet reflecting the systemic risk contribution arising from the interconnectedness of financial institutions (described theoretically in section 2.1.2). This may or may not be a problem, but one interpretation of the results presented in section 4.2 is that SRISK fails to capture some facet of systemic risk, and that may well be interconnectedness.

3.2.2 Definition

I now present the formal definition of the SRISK measure. This extends the capital shortfall framework presented in section 3.1, and follows Brownlees and

Engle. Using (15), SRISK for a firm i at time t is defined as

$$\begin{aligned} \mathrm{SRISK}_{it} &= \mathbb{E}_t(\mathrm{CS}_{it+h}|R_{mt+1:t+h} < C) \\ &= k\mathbb{E}_t(d_{it+h}|R_{mt+1:t+h} < C) - (1-k)\mathbb{E}_t(w_{it+h}|R_{mt+1:t+h} < C). \end{aligned} \tag{20}$$

It is assumed that debt can't be renegotiated in a crisis so its book value stays constant:

$$\mathbb{E}_t(d_{it+h}|R_{mt+1:t+h} < C) = d_{it+h}. \tag{21}$$

It follows from equations (16), (20), and (21) that

$$\begin{aligned} \mathrm{SRISK}_{it} &= kd_{it} - (1-k)w_{it}(1 - \mathrm{LRMES}_{it}) \\ &= w_{it}(k\mathrm{LVG}_{it} + (1-k)\mathrm{LRMES}_{it} - 1), \end{aligned} \tag{22}$$

where LVG_{it} denotes the quasi leverage ratio $\frac{d_{it}+w_{it}}{w_{it}}$. The estimation of LRMES is explained in section 3.2.4.

The SRISK measure can thus be understood as the expected capital shortfall a firm would experience in the event of a systemic crisis. Even though the intuition is fairly simple, the measurement is complicated by the estimation of LRMES, as explained in section 3.1. SRISK can also be defined as an aggregate measure for the financial sector (all N institutions) as follows:

$$\mathrm{SRISK}_t = \sum_{i=1}^{N} (\mathrm{SRISK}_{it})_+, \tag{23}$$

where x_+ denotes $\max\{x, 0\}$. The aggregate SRISK can be thought of as the capital needed for the government to bail out the financial system in the event of a systemic crisis.

For a single institution, a nonpositive SRISK means that in the event of crisis, the firm would still have enough capital to cover its prudential capital requirement. The negative contributions are ignored in the calculation of the aggregate SRISK as it is assumed that the "extra" capital in well-capitalized firms could not be used to cover the shortfall of other firms. Note that this definition addresses the criticism Chen et al. had regarding the "portfolio approach" in the SES measure, as described in section 2.3.3.

SRISK can also be presented in a percentage form, as a systemic risk share of a particular institution:

$$\mathrm{SRISK\%}_{it} = \frac{(\mathrm{SRISK}_{it})_+}{\mathrm{SRISK}_t}. \tag{24}$$

While the exposition by Acharya et al. (2017), presented in section 2.2.1, does not directly apply to SRISK, it still illustrates the kind of framework that makes also SRISK intuitive and comprehensible. In that framework, SRISK roughly corresponds to the equation (13). A part of the regulator's problem is thus to minimize SRISK, or the expected cost of a bailout in a systemic crisis. Note that SRISK does not comment on the regulator's problem any further, e.g. regarding the actual social cost of obtaining the bailout funds (shadow cost of public funds), while it may be included in a framework like the one in section 2.2.1.

SRISK can be also applied as a stress test, as was done by Acharya et al. (2014). This approach allows comparing capital requirements calculated with SRISK to those calculated by regulators. They characterized the SRISK measure as a "mark-to-market" alternative to the regulatory tests that require confidential data. The stress scenario of the SRISK test is determined through the discretionary parameters described in section 3.2.3.

3.2.3 Parameter Specification

Following equation (22), the variables and parameters that determine the value of the SRISK measure for an institution can be illustrated as follows:

$$\mathrm{SRISK}_{it} = f(\mathrm{LRMES}_{it}, \mathrm{LVG}_{it}, w_{it}, k) \tag{25}$$

In other words, SRISK is a function of

- LRMES_{it} — the long-run marginal expected shortfall of the institution, its estimation from financial market data is explained in section 3.2.4
- LVG_{it} — the institution's leverage, determined from its balance sheet
- w_{it} — the value of the institution's equity as determined by the equity market
- k — the prudential capital ratio, determined by regulatory standards or management practice

Also, two discretionary items affect LRMES as described in equation (16):

- C — the equity market decline that characterizes a systemic crisis
- h — the time horizon for C

It is thus obvious that while the first three inputs to SRISK are determined by figures that can be observed objectively, the parameters k, C, and h must be determined by the regulator or assumed by the researcher in a non-regulatory context. These parameters introduce discretion to the SRISK estimation, as does the choice of the LRMES estimation method. From discretion, uncertainty follows.

As emphasized in the paper by Hansen, an important part of evaluating systemic risk measures is dealing with uncertainty. The possibility of choosing "wrong" inputs should not be a reason to abstain from exercises in systemic risk quantification, but the sensitivity of the measure to varying specifications should be emphasized.

In their seminal paper on SRISK, Brownlees and Engle conduct a sensitivity analysis regarding the prudential capital ratio k and the equity market decline C characterizing a systemic crisis. From equation (20), and also intuitively, it is evident that SRISK increases with k and decreases with C.[8] Brownlees and Engle study a baseline specification ($k = 8\%, C = -10\%, h = 22\,(\text{days})$) against two alternative specifications: $k = 10\%, C = -10\%, h = 22$, and $k = 8\%, C = -20\%, h = 22$. They find that the Spearman rank correlation[9] between the baseline and each of the alternatives is high and above 0.90 for the majority of the cross-sections. However, especially institutions in the low end of the rankings exhibit quite high sensitivity to the changes, while rankings for those that are on the top and thus matter the most are relatively stable.

Of course, choices of the parameter values should have reasonable justifications. In the paper by Brownlees and Engle, the value $k = 8\%$ is justified by a back-of-the-envelope estimation of the capital ratio of well-managed banks in normal times. This value is also used as a default in the V-Lab (The Volatility Institute, 2017) data, a project initiated by the same authors. It is a source used in many systemic risk studies, so the use of $k = 8\%$ could be argued to constitute a kind of standard. However, V-Lab uses $k = 5.5\%$ for European institutions due to "differences in accounting standards".

[8] Note that C, as an equity market decline, is expressed here, as in the paper by Brownlees and Engle, as a negative value unlike LRMES, which also denotes a decline but is expressed as a positive value. This is just a matter of convention which should be kept in mind in order to avoid unnecessary confusion. However, it is not consistently followed in the literature.

[9] For an explanation of the metric, see section 3.4.

Brownlees and Engle argue that h should be “sufficiently long” and C “sufficiently extreme”. This leaves much room for interpretation, but the general idea is that the parameters should reflect the characteristics of a systemic crisis. The observations from previous crises may give guidance for the estimation. Brownlees and Engle use $C = -10\%$ and $h = 22$ (days) in their paper as the default specification, V-Lab uses $C = -40\%, h = 6$ (months), and Benoit et al. use $C = -2.52\%, h = 1$ (day), for example. In a real-world setting, the regulator should take the expected cost of the systemic risk externality into account when choosing the discretionary parameters of SRISK.

3.2.4 Estimating LRMES

There are at least three principal methods of estimating LRMES, all presented in the paper by Brownlees and Engle. I summarize these methods in this section.

The method Brownlees and Engle use as the default in their paper is a simulation method incorporating a GARCH-DCC model (Engle, 2002, 2009). The algorithm is appended to their paper, and I will not go through it here. For short, the approach involves specifying time-varying volatility and correlation dynamics of an institution’s stock and the equity market, which are estimated from historical financial market data. Using the estimated parameters, a Monte Carlo simulation of the institution’s equity returns, conditional on a systemic event $\{R_{mt+1:t+h} < C\}$, is performed. This model can incorporate many of the stylized facts in the data.

An alternative to this method is a time-varying copula model described by Patton (2006). In this approach, a copula function is specified to link the institution’s equity returns with the market. It includes a parameter that can be specified to determine the dependence of the two in the lower tail of returns, which is a specific merit in it that Brownlees and Engle point out.

However, the method I focus on here is one that Brownlees and Engle call the static bivariate normal model. This approach assumes that the firm’s and the market’s logarithmic equity returns are independent and identically distributed from a bivariate normal distribution with zero mean. The estimation requires calculating the market volatility, the firm volatility, and correlation parameters σ_m, σ_i, and ρ_{im}, respectively, from historical financial market data.

Using these parameters, the CAPM beta for the firm's stock is as follows:

$$\beta_i = \rho_{im} \frac{\sigma_i}{\sigma_m}. \tag{26}$$

In this model, LRMES can be approximated[10] as follows:

$$\text{LRMES}_{it} = -\sqrt{h} \beta_i \mathbb{E}(r_{mt+1} | r_{mt+1} < c), \tag{27}$$

where

$$\mathbb{E}(r_{mt+1} | r_{mt+1} < c) = -\sigma_m \frac{\phi\left(\frac{c}{\sigma_m}\right)}{\Phi\left(\frac{c}{\sigma_m}\right)},$$

where $\phi(x)$ is the standard normal probability density function and

$$c = \frac{\log(1 + C)}{\sqrt{h}}.$$

However, it is further simplified in the V-Lab environment as

$$\text{LRMES}_{it} = 1 - \exp(\log(1 - C) \times \beta), \tag{28}$$

which I shall refer to as static LRMES from now on. This definition assumes a constant $h = 6$ months.

According to the V-Lab documentation, this formulation allows for "transparency, reproducibility, and flexibility", and it makes the maths here a bit easier to understand. In section 4, I present the results of my analysis on whether this formulation that connects SRISK to systematic risk in quite a direct way still allows making the distinction between systematic risk and systemic risk and producing otherwise meaningful results.

A variant of the static LRMES, with $h = 6$ months and $C = -40\%$, is

[10] LRMES has the following exact closed-form solution:

$$\text{LRMES}_{it} = -\exp\left(\frac{h}{2}(\beta^2 \sigma_m^2 + (1 - \rho^2)\sigma_i^2)\right) \frac{\Phi\left(\frac{\beta \log(1+C) - h\beta^2\sigma_m^2}{\sqrt{h}\beta\sigma_m}\right)}{\frac{1}{2} + \frac{1}{2}\operatorname{erf}\left(\frac{\beta \log(1+C)}{\sqrt{2}\sqrt{h}\beta\sigma_m}\right)} + 1,$$

where $\operatorname{erf}(x)$ is the error function and $\Phi(x)$ is the standard normal cumulative distribution function.

presented by Acharya et al. (2012). It defines the LRMES as

$$\text{LRMES}_{it} = 1 - \exp(-18 \times \text{MES}_{it}), \tag{29}$$

where MES_{it} is the one-day expected loss if the market return is less than -2%.

With LRMES defined as in (28), the SRISK illustration in (25) can be reformulated as

$$\text{SRISK}_{it} = f(\text{LVG}_{it}, \beta_{it}, w_{it}, k, C). \tag{30}$$

Brownlees and Engle present a comparison of the three methods for calculating SRISK. They compare them pairwise using a Granger causality test, based on a predictive regression

$$\Delta \log \text{SRISK}^A_{t+1} = \alpha_0 + \alpha \Delta \log \text{SRISK}^A_t + \beta \Delta \log \text{SRISK}^B_t + u_t, \tag{31}$$

where $\Delta \log \text{SRISK}^A_t$ and $\Delta \log \text{SRISK}^B_t$ denote changes in log aggregate SRISK based on alternative LRMES estimators A and B, respectively, and u_t is the error term.

LRMES based on estimator B is said to lead the LRMES based on estimator A if the regression coefficient β is significantly different from zero. The intuition is that if B leads A, then B will provide information about the evolution of the SRISK measure before A. If an estimator is not lead by any other estimator, Brownlees and Engle call it a timely estimator. They perform this regression for each pair of the three estimators, and find that the simulation-based (GARCH-DCC) estimator of LRMES (I will call it simulated LRMES from now on) is not lead by any other estimator, but leads them. The static LRMES is found to be lead by the others and not to lead either one of the other measures. They also report a rank correlation analysis similar to the one conducted with the k and C parameters, and find that the correlations are quite high, ranging from 0.74 to 1.00, but that the differences between the LRMES estimators are particularly pronounced during the crisis period.

Brownlees and Engle conclude that the static LRMES does not provide a timely measure of SRISK. They also note that the specification (27) has a negative approximation error and should only be used for a short horizon. They also conclude that the simulated LRMES is a particularly useful estimator. The conclusion that static LRMES has shortcomings is supported by the results presented in section 4.

3.3 Data

The main source of the data in the empirical section of this thesis is the V-Lab (2017) website of The Volatility Institute of New York University Stern School of Business. Created under the direction of Nobel Laureate, Professor Robert F. Engle, the service reports various measures related to volatility, liquidity, and other important market variables. It also publishes data on systemic risk, and is updated daily.

The systemic risk data, which is focused on the SRISK measure, is composed in accordance to the SRISK framework I have presented in section 3.2. There are three datasets: one for U.S. financial institutions using the static LRMES estimator, one for global (non-U.S.) financial institutions, also with static LRMES, and one for U.S. financials with the simulated LRMES.

In my empirical analysis, I study specifically the static LRMES data for U.S. financials. There is a handful of reasons for this choice. An important reason is that the sheer variety of the proposed estimators for LRMES presents a challenge for researchers and regulators. Different methods of estimating LRMES coexist in the literature, yet there seems to be very limited literature on the merits and shortcomings of these methods. Knowing which LRMES estimator to use, and why, would be very relevant information for anybody studying SRISK, especially considering the differences between the estimators pointed out by Brownlees and Engle.

It is reasonable to assume that the V-Lab data holds a special position as a data source for the SRISK literature. Not only is the project initiated by the people behind SRISK itself, but it also provides an extensive dataset that caters well to the needs of systemic risk research. While it is difficult to exactly measure how many studies have taken their data from the V-Lab service, it is easy to find several such papers. However, it seems that the limitations regarding the LRMES estimator are not widely considered, or at least not reported in the papers using the V-Lab data. One paper (Pierret, 2015) chooses the static LRMES corresponding to (29) for U.S. banks without any reported justification, while the simulated LRMES would have been available. Some papers, such as one by Kim et al. (2016), doesn't even report which definition of LRMES, and thus SRISK, they use. Also, specifically for non-U.S. studies, the choice of using V-Lab data means choosing to use the static LRMES. For an example of these see e.g. Nucera et al. (2016). The situation is further complicated by the fact that the formula for the static LRMES was changed in V-Lab from (29) to (28) at some point in time. The specification (28) itself has not been discussed in any theoretical literature, but only in the V-Lab documentation, where it is justified with mainly practical aspects.

The choice of LRMES estimation method is particularly important when considered from the regulatory perspective. Suppose that based on favorable analyses of the SRISK measure in the literature, the regulator chose SRISK to form the basis for a tax or a capital requirement. Now, as a real-world example from the V-Lab data I study, the SRISK calculated for J.P.Morgan on 31 December 2015 with the static LRMES is USD 53.7 billion. This denotes the amount of additional capital the firm would need if a systemic crisis hit, and in order to mitigate the effect for the real economy stemming from this expected capital shortfall, the regulator would impose a tax or a capital requirement that would take into account the magnitude of the expected shortfall. However, the corresponding figure calculated with the simulated LRMES is 75.7 billion, 22 billion dollars higher! This example illustrates that the differences between different LRMES estimators should be studied carefully if SRISK was to be promoted as a viable measure of systemic risk for regulatory purposes.

By studying the data composed with the static LRMES estimator, I hope to shed light on its viability as a basis for SRISK for regulatory purposes. I also hope to contribute to the overall understanding of the SRISK measure.

The main dataset I use is based on the Systemic Risk Analysis: U.S. financials dataset with the static LRMES estimator, provided on the V-Lab website. I have extracted monthly data for the period December 2001 – January 2017, or 182 time periods. The data corresponds to the last business day of the month, and is estimated recursively, i.e. using only data available at the specific date. The dataset includes figures for 97 major U.S. financial firms, but some of the firms don't have observations for the whole timeframe, as they may have gone bankrupt or have been founded after 2001. In total, there are 14837 observations of each variable.

The variables collected and calculated by V-Lab include the following:

- SRISK — as defined in (22) with $k = 8\%$, $h = 6$ months, and $C = -40\%$, in millions of U.S. dollars (MUSD).

- SRISK% — as defined in (24), in percent.

- LRMES — as defined in (28), in percent.

- Beta — as defined in (26), where the market corresponds to the S&P 500 index, a widely followed benchmark of approximately 500 U.S. companies with a high market capitalization (total equity value).

- Cor — dynamic conditional correlation between the equity return on an institution's stock and the return on the S&P 500 index, estimated with a DCC model.[11]
- Vol — annualized volatility of an institution's stock, estimated with a GJR-GARCH model (Engle, 2002, 2009), in percent.
- Lvg — as defined in section 3.2.2.

Summary statistics for some of the variables are presented in table 1.

Statistic	Mean	St. Dev.	Min	Max
SRISK, MUSD	1027	21617	−205140	161421
LRMES, %	44.5	13.2	−21.4	100
Beta	1.226	0.638	−0.38	31.4
Cor	0.594	0.145	−0.13	0.92
Vol, %	34.5	37.2	0	2029
Lvg	8.2	38.3	1	4106

Table 1: Summary statistics for the main dataset. N=14837 for each variable.

I have also calculated the aggregate SRISK, using (23). It is presented in figure 1.

There are three reasons for the choice of U.S. data instead of the global dataset. First, using the U.S. data allows making a comparison between the static and simulated LRMES estimators. Second, for example one of the most important European banks, Crédit Agricole, has a complex organizational structure with both listed and cooperative entities that may affect the reliability of the SRISK figures calculated from the data of the main holding company. Also, in the V-Lab data, some of the local banks that are both owners and clients of the cooperative are reported separately. The third reason for using U.S. data is that I find it reasonable to assume that the U.S.–based team has put less effort into dealing with this kind of idiosyncracies in the European dataset than into composing the U.S. one.

[11] I don't consider it necessary to discuss this in more detail as the variable is not directly used in my analysis.

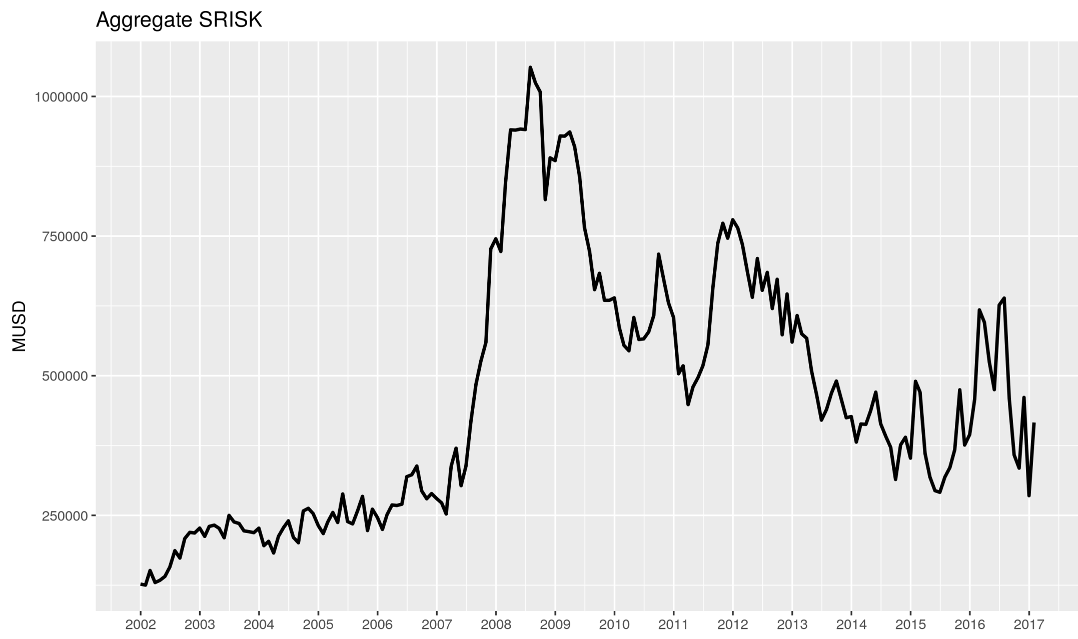

Figure 1: Aggregate SRISK in the financial system over time.

For the purpose of making comparisons, I extend my data with alternative indicators of systemic risk. I use the Financial Stability Board's 2016 list of global systemically important banks (FSB, 2016) as one point of comparison. The data contains a list of 30 financial institutions from North America, Europe, and Asia, that are categorized into four buckets according to the BCBS (2013) scoring methodology briefly presented in section 2.3.3. According to the BCBS, the scores are supposed to reflect "the size of banks, their interconnectedness, the lack of readily available substitutes or financial institution infrastructure for the services they provide, their global (cross-jurisdictional) activity and their complexity". This classification, based on data gathered on 31 December 2015, is presented for U.S. firms in table 3.

I also use the V-Lab data for the 31 December 2015 cross-section for non-U.S. institutions and U.S. institutions with simulated LRMES in the comparisons. The variables are the same as in the main dataset with the slight differences discussed above.

3.4 Estimation

The purpose of this study is to find out if the SRISK measure with a static LRMES estimator could serve as a useful measure of systemic risk. The usefulness of a particular measure can be determined in many ways as described in section 2.4, but the approach I take is that I try to find out whether the SRISK measure could provide information not available from a purely market-based measure (CAPM beta), and if the SRISK estimates of systemic risk correspond to the regulatory classification.

I would consider SRISK probably useful if I found that it both

1. provides rankings that differ from CAPM beta by emphasizing institutions that reflect the characteristics of systemic risk (e.g. size, leverage, interconnectedness), and
2. provides rankings that roughly correspond to those provided by the regulators.

It would also have to be able to provide this without being totally explained by one of its constituents (e.g. leverage), in which case only the constituent indicator could be considered a probably useful measure.

The finding that the SRISK measure is probably useful would not be a total guarantee of its usefulness, and further research would still be required. On the other hand, the finding that SRISK either has almost one-to-one correspondence with the CAPM beta, or that its predictions are totally different from the ones provided by the regulators, would cast a serious doubt on its usefulness and serve as evidence of SRISK with static LRMES being a suboptimal estimator of systemic risk.

Having an estimator that can reliably identify the institutions that are most important systemically would be very beneficial for the regulator, and correspondence to the current regulatory rankings would indicate that SRISK could work as a measure of systemic risk. However, it is a completely different question whether the absolute level of SRISK gives useful information for the regulator. As it is virtually impossible to determine ex ante, researchers need to conduct analyses of past capital injections like the one presented by Brownlees and Engle, for example. I limit my analysis to the rankings, as I consider that an intuitive starting point for the evaluation.

Benoit et al. conduct a simple analysis where they compare the systemic risk rankings of institutions provided by the CAPM beta and a variant of the MES measure. They find that the two are highly correlated, and they consider that worrisome for a number of reasons. If the correlation is very high, the same result that could be achieved with MES could be achieved with beta in a more simple way. However, it is known that beta is not a good measure of systemic risk, as it doesn't capture many of the important sources of systemic risk such as the size of the institution. This diminishes the credibility of MES as a systemic risk measure, too. Another reasons include the introduction of confusion between systematic and systemic risk, and procyclicality. In the latest financial crisis, for example, betas increased notably and a measure that is found to be highly correlated with betas would be inherently procyclical, and if used as a basis for a tax, it could further aggravate the crisis. I present a similar analysis for SRISK and the CAPM beta.

Brownlees and Engle present a ranking comparison between SRISK and a selection of alternative indicators using the Spearman rank correlation coefficient (Spearman, 1904). The Spearman rank correlation coefficient (or Spearman's ρ) for two variables is calculated by ranking the variables and calculating the standard Pearson correlation coefficient (Pearson, 1895) for the ranks. It is a common non-parametric measure of correlation, and, as opposed to the Pearson's ρ, can capture non-linear correlation as well. I compare the rankings for financial institutions according to alternative measures with the rankings provided by SRISK using Spearman's ρ.

Another non-parametric correlation measure is the Kendall rank correlation coefficient or Kendall's τ (Kendall, 1938). Its calculation involves ordering the two sets into pairs and counting the different pairs between the two sets of pairs. The resulting coefficient takes values from -1 to 1, with 1 denoting identical orderings, -1 totally reverse orderings, and 0 that the sets are independent. It can be interpreted in a probabilistic way as the probability difference between the two sets being actually in the same order, or

$$\tau = P(\text{same}) - P(\text{different}).$$

For a more detailed explanation, see e.g. Abdi (2007).

In order to demonstrate the linear relationships between different variables, I also report some OLS regressions. All the analysis is conducted using the R statistical language (R Core Team, 2016) with the help of the RStudio environment (RStudio Team, 2016). I illustrate the analysis with graphs and figures I have drawn using the ggplot2 package for R (Wickham, 2009).

4 Results

In this section, I present the results of my analysis.

4.1 Systemic vs. Systematic

In order to be a useful systemic risk measure, SRISK should be able to provide some information not available from using just a systematic risk measure, such as the CAPM beta. In the paper by Benoit et al., high linear correlation between the CAPM beta and the MES measure is found to severely restrict the usefulness of MES as a systemic risk measure, and I present here a similar analysis for SRISK. An OLS regression of the long-term arithmetic mean of SRISK on the long-term arithmetic mean of the CAPM beta across the 97 institutions is reported in table 2 and illustrated in figure 2.

There is a significant and positive relationship between the CAPM beta and SRISK, which is not surprising as beta is a component of SRISK through the LRMES. However, the coefficient of determination (R^2) is quite low, 0.17, which suggests that SRISK may be able to incorporate more information than just the systematic risk component of systemic risk. The illustration in figure 2 supports this point of view.

	Dependent variable:
	SRISK, MUSD
CAPM beta	21633^{***}
	(4928)
Constant	-25016^{***}
	(6216)
Observations	97
R^2	0.169
Adjusted R^2	0.160
Residual Std. Error	14243 (df = 95)
F Statistic	19.274^{***} (df = 1; 95)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 2: Results from the OLS regression of SRISK on CAPM beta, both taking their long-term average over the whole time-series in the main dataset. Standard errors in parentheses.

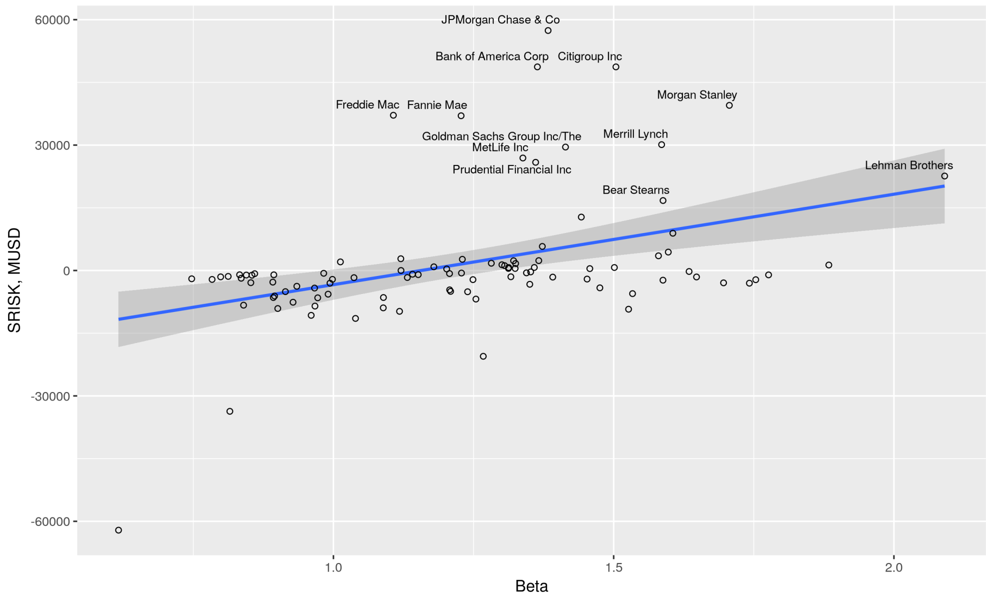

Figure 2: CAPM beta and SRISK, average cross-section. The line corresponds to the regression reported in table 2, and the shaded area is the 95% confidence region.

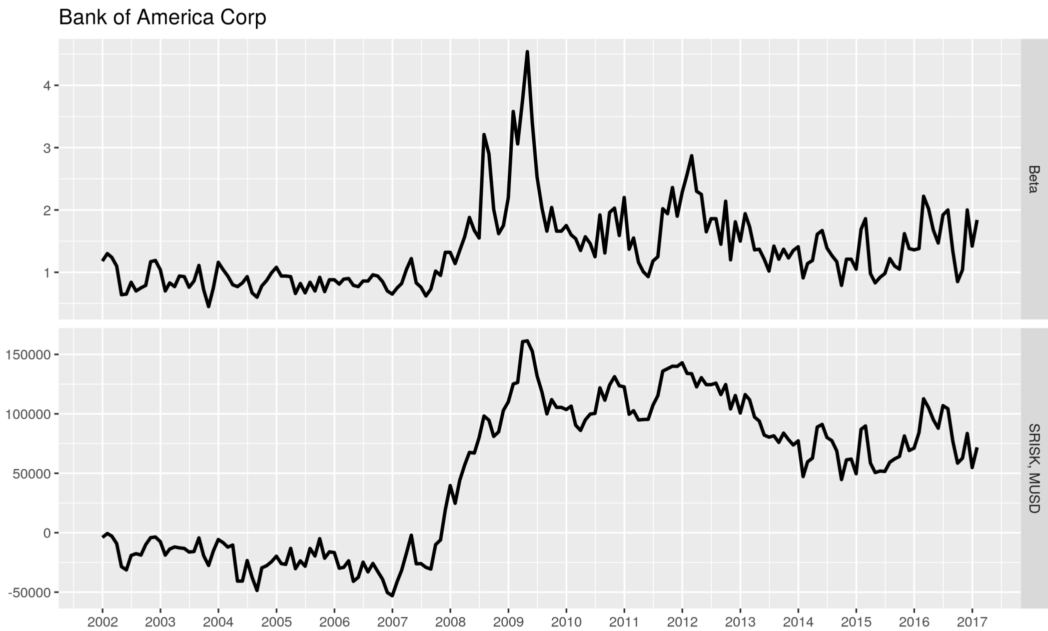

Figure 3: CAPM beta and SRISK, time series for Bank of America.

Even though the connection between the CAPM beta and SRISK is significant in the OLS regression, beta is unable to predict a significant number of the observations, including the ones with the highest average SRISK. Of the 12 institutions with the highest SRISK (and thus systemic importance), only one

(Lehman Brothers) is within the 95% confidence region for the estimator, as seen in figure 2. A reader familiar with the history of the 2007–2009 financial crisis will recognize several names of failed institutions among the ones with the highest SRISK measures.

Benoit et al. show that there is a proportional relationship between the ΔCoVaR measure and an entirely market-based measure of tail risk, VaR. This supports the notion that the variation in ΔCoVaR across the time series dimension can be entirely captured by VaR, similarly as the variation in MES is found to be captured by the CAPM beta across the cross-section. An an illustrative example, figure 3 presents the time series of both CAPM beta and SRISK for Bank of America, showing that the relationship between beta and SRISK is not constant across the time-series. The case is similar for other institutions in the sample as well.

While not proving that SRISK is a good measure of systemic risk, these cross-sectional and time-series characteristics demonstrate that it may be able to evade the criticism presented by Benoit et al. regarding some alternative measures of systemic risk. These preliminary results thus suggest that SRISK might prove useful, and are supportive of further analysis.

4.2 Comparisons with Alternative Measures

I now compare the rankings provided by the SRISK measure to possible alternatives using the cross-sectional data from 31 December 2015. The chosen date corresponds to the date of the latest FSB G-SIB data.

Figure 4 corresponds to figure 2, but for the single cross section on 31 December 2015. Some specific institutions that rank high according to beta, SRISK, and the G-SIB list are named. The overall picture of the relationship is very similar to figure 2.

Table 3 presents rankings of some of the most important U.S. financial institutions on 31 December 2015 according to alternative measures of systemic risk. Note that while beta is not a systemic risk measure, it is included for comparison. The table includes top 10 institutions according to the two different variants of SRISK based on the LRMES estimation method and beta, as well as all U.S. institutions featured in the FSB list of global systematically important banks. The G-SIB list is presented in categories, or buckets, where a higher bucket number denotes higher systemic importance and risk. From the table and figure 4, it is apparent that ranking institutions on betas provides very different results than ranking them on either version of SRISK. In this cross-section, the two versions of SRISK create fairly similar rankings, and the higher G-SIB buckets roughly correspond to higher SRISK rankings. Note that the G-SIB list only includes banks.

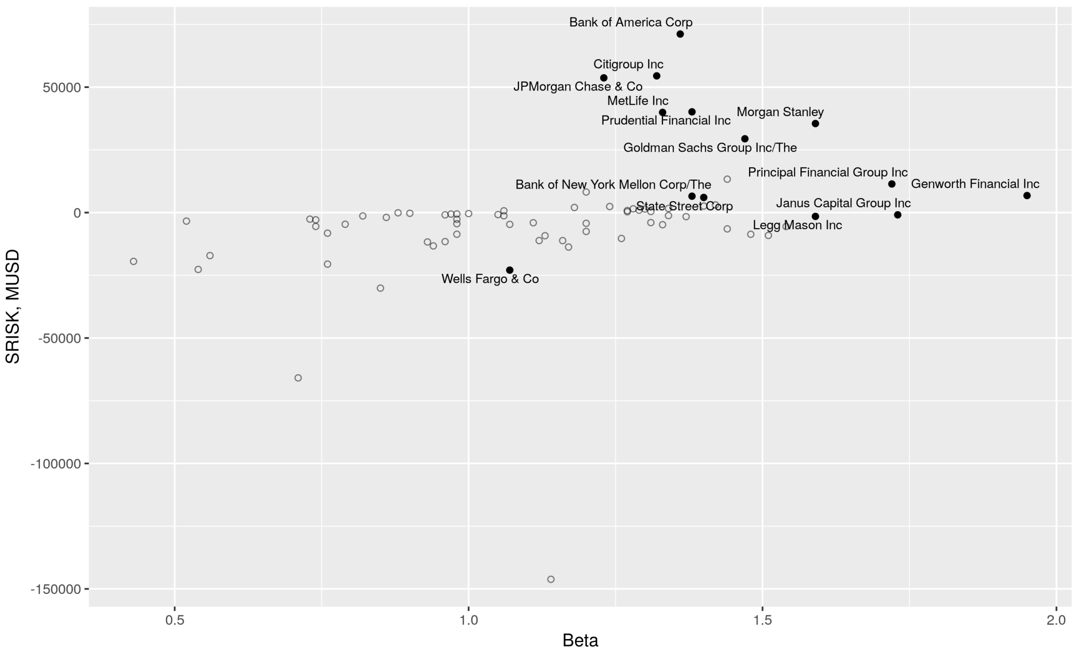

Figure 4: CAPM beta and SRISK as of 31 December 2015.

The possible relationship between SRISK and the G-SIB buckets warrants further consideration. For this analysis, I use a dataset that is combined of the main dataset, the global financials dataset from V-Lab and the G-SIB list. The G-SIB list includes 30 institutions, but I exclude Groupe BPCE as it is not included in the V-Lab dataset.[12]

Figure 5 illustrates the relationship between the G-SIB buckets and SRISK figures, and the corresponding OLS regression is reported in table 4. SRISK is regressed on the G-SIB buckets, and the intuition is that if the G-SIB buckets are assumed to be a reliable indicator of systemic risk, a higher bucket should translate into a higher SRISK, and the beta coefficient of the regression would determine how big an increase in SRISK corresponds to a step to a higher G-SIB bucket.

[12]Its subsidiary Natixis is.

Institution	SRISK ranking (static)	SRISK ranking (simu-lated)	Beta ranking	FSB G-SIB bucket
Bank of America	1	2	18	3
Citigroup	2	3	23	4
J.P.Morgan	3	1	33	4
Prudential	4	5	15	—
MetLife	5	4	21	—
Morgan Stanley	6	6	4	1
Goldman Sachs	7	7	9	2
Lincoln Financial Group	8	8	10	—
Principal Financial Group	9	10	3	—
The Hartford	10	11	34	—
Genworth	11	13	1	—
BNY Mellon	12	12	16	1
State Street	13	14	13	1
Capital One	23	9	46	—
Janus Capital Group	32	36	2	—
Legg Mason	37	39	5	—
Charles Schwab	52	51	6	—
BlackRock	57	37	8	—
Franklin Resources	58	61	7	—
Wells Fargo	71	23	45	2

Table 3: Systemic risk rankings for a selection of U.S. financial institutions on 31 December 2015 according to various measures. The main dataset includes 74 institutions. Note that a higher bucket number indicates higher systemic importance, as opposed to the rankings, where a smaller number indicates a higher ranking.

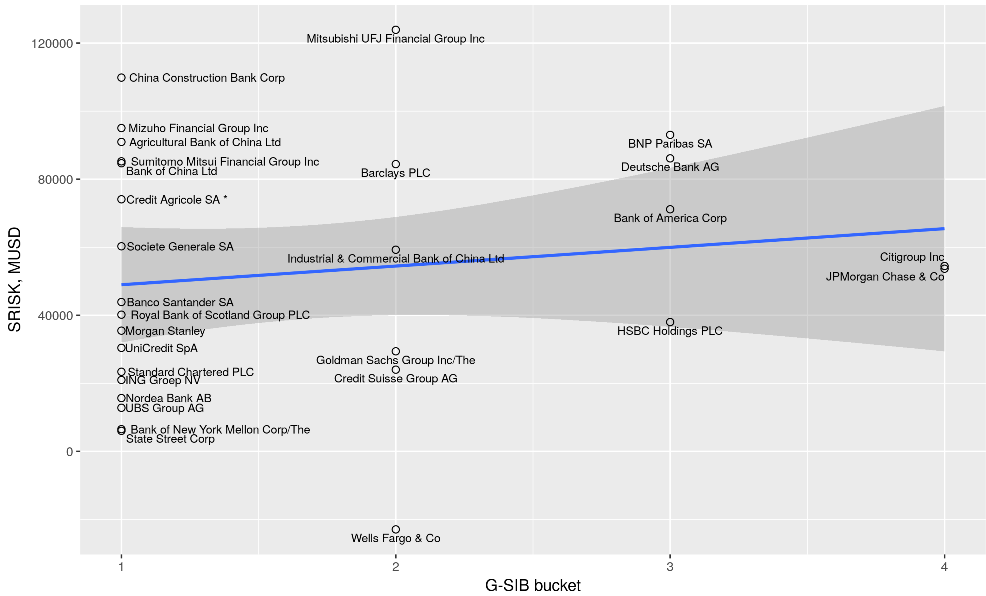

Figure 5: The relationship between SRISK figures and the G-SIB buckets, using 31 December 2015 data.

	Dependent variable:
	SRISK, MUSD
G-SIB bucket	5495
	(7034)
Constant	43489***
	(13637)
Observations	29
R^2	0.022
Adjusted R^2	−0.014
Residual Std. Error	36008 (df = 27)
F Statistic	0.610 (df = 1; 27)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 4: Regression of SRISK on the G-SIB buckets.

The results from the regression, reported in table 4, show that there is no significant relationship between the G-SIB buckets and the SRISK figures for the 29 institutions included in the combined dataset. The coefficient of determination is very low and the Pearson correlation between the two variables is only 0.15 and insignificant, having a p-value of 0.44.

Many of the ranks are tied in the G-SIB data because of the bucketing approach, but the Kendall's τ can be used as an alternative to the Spearman's ρ in such a case (see e.g. Field et al., 2012, p. 225). The estimated Kendall's τ between the G-SIB buckets and SRISK is only 0.13, with a p-value of 0.40, so there is no significant correlation. At least in this sample, it thus seems that either SRISK or the G-SIB list is unable to identify the variation in the systemic importance of some of the possibly most systemically important institutions. Furthermore, if the G-SIB buckets are believed to be able to capture the cross-sectional differences in systemic risk, it must be concluded that at least here, SRISK is not. That would call its usefulness into serious question.

I now turn to comparing the SRISK measure with the other measures reported in the V-Lab dataset that serve as its constituents. The approach I take here is the same as presented by Brownlees and Engle.

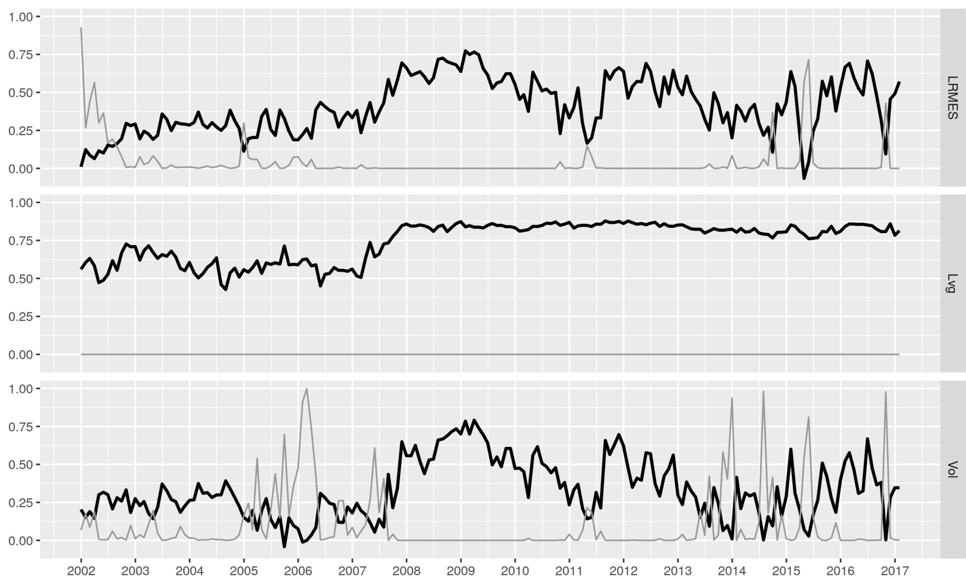

Figure 6: Rank correlation over time between SRISK and three different risk indicators: LRMES, firm leverage, and volatility. Black line denotes the Spearman correlation coefficient and gray line is the p-value. Except for periods with a very low correlation, the value of the correlation coefficient between SRISK and the risk indicator is statistically significant at 0.01 significance level.

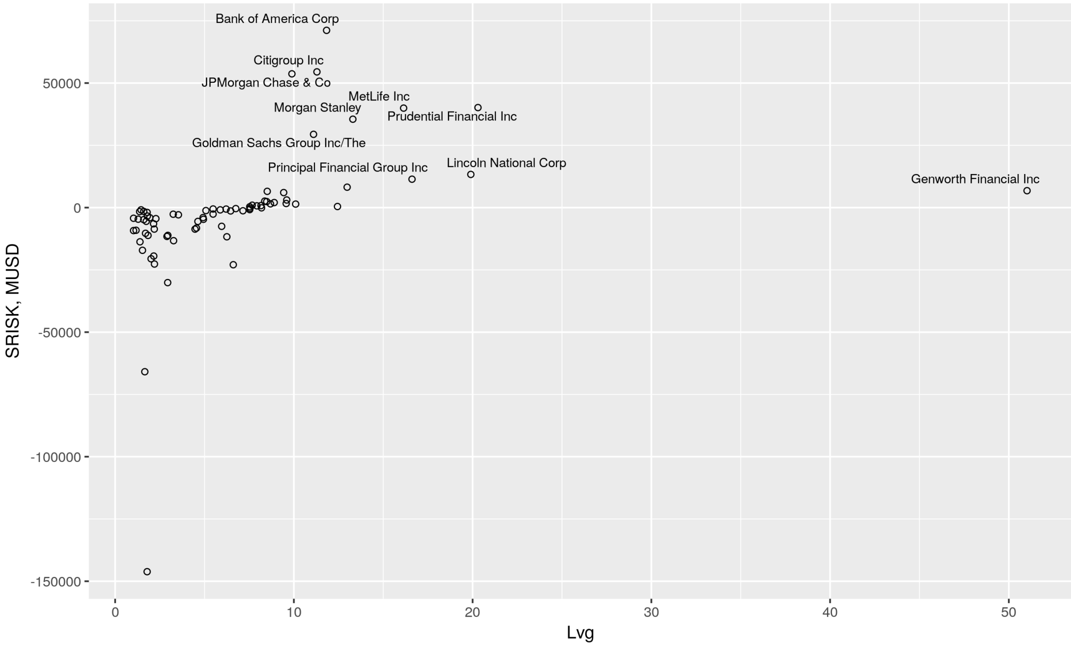

Figure 7: SRISK plotted against leverage on 31 December 2015: many institutions rank high on both measures, but not necessarily in the same order.

Figure 6 presents the Spearman's ρ between SRISK and LRMES, leverage, and the volatility of the firm's stock over time. The correlation coefficient is calculated separately for each time period. As the Spearman's ρ considers ranks, the results for LRMES and beta are equivalent in accordance to (28). The time series for the correlations with their respective p-values show that the significance and magnitude of the correlation vary quite much across time and the variables. The correlation between SRISK and leverage is quite high and $p < 0.01$ for the whole timeframe, while in the correlations with LRMES and with the equity volatility there is more variation. A notable feature is that all correlations rise at the onset of the 2007–2009 financial crisis.

The reported correlations are higher and more significant than those reported by Brownlees and Engle for the SRISK with simulated LRMES and the same three variables. It is also very interesting to note that the figures they report for SRISK–leverage and SRISK–LRMES correlations decrease towards zero at the onset of the crisis and never exceed 0.40. The SRISK–volatility correlation is almost never significant and hovering around zero in their study.

Figure 7 presents the cross-sectional scatterplot of SRISK and leverage on 31 December 2015. The Spearman correlation is 0.81, and many of the highest-ranking institutions are the same when measured with leverage and SRISK. Even though they are not in the same order, the high correlation is a worrisome finding in the light of the criteria for the probable usefulness of SRISK presented in section 3.4.

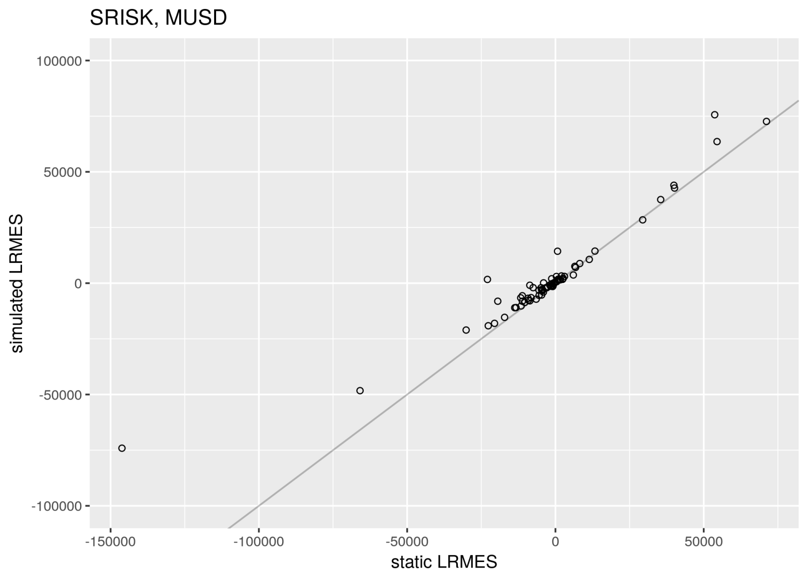

Figure 8: SRISK with simulated LRMES plotted against SRISK with static LRMES on 31 December 2015.

Figure 8 presents a comparison of the SRISK measure estimated with the two alternative LRMES estimators. It illustrates the negative approximation error in the static LRMES estimator noted by Brownlees and Engle. The Spearman correlation coefficient is 0.94, and the grey line denotes the equality between the two measures. It is apparent that the SRISK approximated with the static LRMES is a biased estimator of the simulated SRISK, giving mostly lower SRISK estimates. For the purposes of regulation, too low estimates may be harmful if they leave some systemic risk effects being uncaptured.

The figure 9 and the table 5 report the relationship between the G-SIB buckets and the U.S. institutions for which SRISK with simulated LRMES is available. The Pearson correlation between the two variables is 0.82, having a p-value of 0.012, but any inferences about the relationship should be very cautious as the sample is very small. Also, the corresponding relationship with static LRMES and only U.S. institutions (not reported) doesn't look very different.

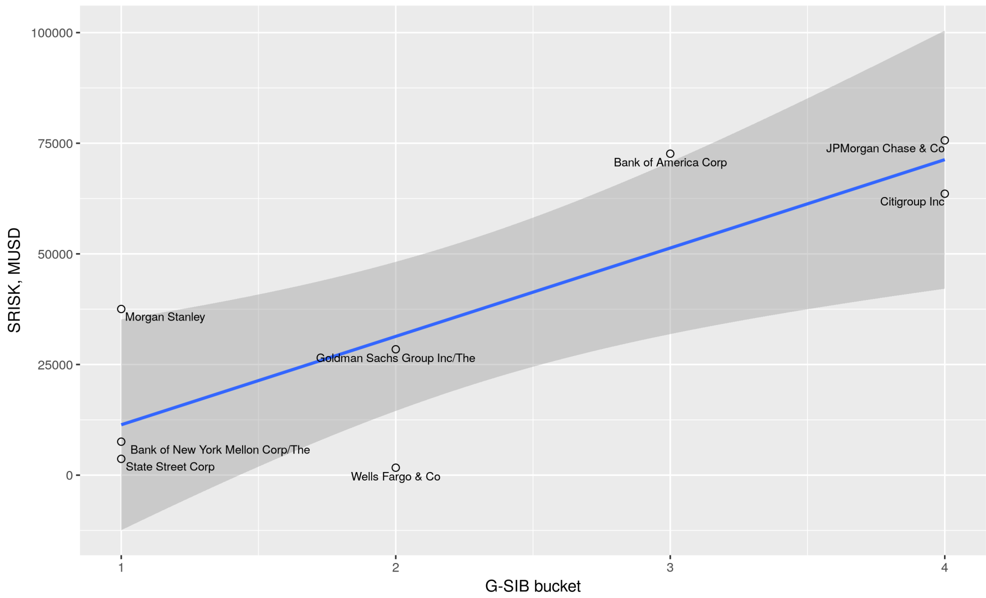

Figure 9: The relationship between SRISK figures with simulated LRMES and the G-SIB buckets, using 31 December 2015 data.

	Dependent variable:
	SRISK, MUSD
G-SIB bucket	19970**
	(5621)
Constant	−8584
	(14330)
Observations	8
R^2	0.678
Adjusted R^2	0.624
Residual Std. Error	19061 (df = 6)
F Statistic	12.624** (df = 1; 6)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 5: Regression of SRISK with simulated LRMES on the G-SIB buckets.

5 Discussion

I now discuss the interpretation of the above results, their significance, and limitations of the study. I also present some thoughts on the shortcomings of the SRISK measure and how it could be improved.

5.1 Interpretation of the Results

The results suggest that there are significant limitations to the usefulness of SRISK with static LRMES as a measure of systemic risk. Even though it is able to differ from the CAPM beta, a purely systematic risk measure, it has a very weak relationship with the regulatory rankings (buckets) of global systemically important banks and exhibits a high correlation with leverage, one of its constituents. The results thus seem to suggest that the rankings provided by SRISK with static LRMES could be quite closely replicated by just using an institution's leverage as a systemic risk measure. However, as discussed in this thesis, there are several driving forces behind systemic risk, of which leverage is only one.

The results can be compared with those obtained by Brownlees and Engle for the SRISK measure with simulated LRMES. They, too, find that the Spearman correlation between SRISK and leverage is significant, yet it is much lower than what is observed here. It seems that their measure coincides with the constituents of SRISK in a different way, even though the formula is the same.

Brownlees and Engle also find that the simulated estimator exhibits lower correlations with its constituents and alternative measures (such as SRISK with static LRMES) during the 2007–2009 crisis period. This is a promising finding compared to the results presented in this work, as rising correlations to rising market measures during crisis introduce harmful procyclicality to the measure.

It is an interesting finding that even though these two variants of SRISK appear superficially to be the same measure, their more careful analysis reveals significant differences in how they relate to their underlying components. This highlights the importance of careful analysis of the underlying model specification when applying these measures in practice. In addition to their internal mechanisms, the actual results of the shortfall estimations conducted with these alternative SRISK specifications may differ greatly, as evidenced e.g. by figure 8.

Even though there appear to be significant shortcomings in the SRISK measure with the static LRMES estimator, it might still provide some additional information about the systemic risk of institutions when compared with e.g. the CAPM beta. However, all the results suggest that the simulated SRISK is more useful and further research should concentrate on that measure. Researchers on the topic should also be very careful when making inferences based on the static estimator of SRISK, and the justification for the LRMES estimator choice should be clearly discussed. An interesting topic for further research would be to study how sensitive the simulated LRMES itself is to reasonable changes in the underlying asset price simulation model specification.

As illustrated by figure 5 and the corresponding regression reported in table 4, the relationship between the regulatory buckets and the SRISK figures is very weak. This can be interpreted so that SRISK may not be able to capture the variation in the systemic risk contributions across institutions very accurately, even though the institutions in the G-SIB list generally rank in the higher end of the SRISK-based list (see table 3). These results don't confirm the notion that the simulated SRISK corresponds to the regulatory buckets any better than the static one.

Some possible reasons for these differences between SRISK and the regulatory approach can be found by looking at the methodological differences between these models. While the regulatory approach is not trying to estimate the expected capital shortfall of the institutions, the risk of a bank's failure is inherently linked to the lack of capital. It thus seems reasonable to assume that in essence, SRISK and the regulatory buckets are trying to measure the same thing. However, the submeasures included in these two approaches differ somewhat. As seen in (30), SRISK incorporates the leverage (LVG_{it}), size (w_{it}), and the systematic risk of an institution (β_{it}), while the regulatory buckets take into account "the size of banks, their interconnectedness, the lack of readily available substitutes or financial institution infrastructure for the services they provide, their global (cross-jurisdictional) activity and their complexity" (BCBS, 2013). It can be argued that the categories must be at least somewhat overlapping, but not totally.

As discussed in section 3.2.1, in order to capture the portion of systemic risk arising from the interconnectedness of an institution, it should be implicitly included in LVG_{it}, w_{it}, or β_{it}. However, it is not at all evident that any of these variables should vary according to the interconnectedness of an institution, or that two institutions having similar levels of these variables would exhibit a similar level of interconnectedness within the financial system. Alternatively, the interconnectedness should have an impact on the probability of the systemic event $\{R_{mt+1:t+h} < C\}$, and it is evident from the formula for static LRMES (28) that the interconnectedness is not accounted for explicitly. The simulation framework described in more detail in the paper by Brownlees and Engle doesn't consider the interconnectedness of the institutions per se either, but relies on firm-market correlations in modeling the equity returns. It is not entirely implausible that the interconnectedness of the financial institutions would have a similar effect on both the systematic risk and the systemic risk of the institutions, but it would nevertheless be quite a strong assumption that would require an explicit justification.

As discussed in section 2.1.2, the structure of the network of links between financial institutions may have effects on how large the systemic risk externalities are. It is thus possible that the usefulness of SRISK is restricted by its lack of ability to incorporate the interconnectedness, or network, characteristics of the financial system, which are explicitly accounted for in the FSB G-SIB methodology.

The possibility of extending SRISK by incorporating network-based indicators, such as centrality measures, should be considered. One way to account for interconnectedness would be to model the financial network based on the interbank exposures that financial institutions have. While quite possibly suffering from practical difficulties, this kind of an approach might be able to improve the accuracy of SRISK. Abbe et al. (2012) present the possibility of calculating risk measures while maintaining privacy, and such methods might provide a practically feasible way for the regulator to assess the interconnectedness of banks. However, more understanding on the effects of the structure of the interbank network is needed before it can be used to accurately quantify systemic risk in a framework like SRISK.

5.2 Limitations of the Study

There are some limitations to the interpretation and applicability of the results of this study.

An obvious limitation is that comparing the SRISK figures to the regulatory buckets implies that the regulators have perfect knowledge of the level of systemic risk contributions of all these institutions. This is quite unrealistic, given that there is no unified framework of determining how the sources of systemic risk should translate into systemic risk contributions of individual financial institutions. However, they might still give an approximation and incorporate the best data and some of the most advanced understanding of systemic risk that exists at the moment.

Also, no promising result of the usefulness of SRISK based on rankings provided by the SRISK measure should be interpreted as SRISK providing accurate data about the actual expected capital shortfalls. In other words, even if SRISK happened to construct a ranking that would be deemed reasonable, it might be that the absolute levels of SRISK didn't provide almost any useful information.

In the results presented above, some analyses of the 31 December 2015 data were considered to support the notion that SRISK might not be useful. The data is thus limited to only one point in time, and conducting several similar analyses or more detailed comparisons of time series might still give evidence in favor of SRISK. However, I consider it reasonable to expect that SRISK should work at least reasonably well at every single point of time in order to be considered a relevant measure from the regulatory perspective.

6 Conclusion

As an inquiry into systemic risk and its measurement, I have presented an empirical assessment of the SRISK measure of systemic risk, estimated with a static estimator for the long-run marginal expected shortfall (LRMES) component. I now summarize the conclusions of my analysis.

Previous literature has presented conflicting accounts of the usefulness of SRISK, but the promising results obtained by e.g. Brownlees and Engle suggest that the measure should not be abandoned but studied further. The empirical literature, including evaluations, on SRISK is still quite limited and may suffer to some extent from the lack of thorough understanding of the LRMES model specification. The results presented in this study may help to guide further research to acknowledge the limitations of the static LRMES estimator in particular.

In my empirical analysis, I find that the systemic risk evaluations based on the SRISK measure differ from the ones provided by the CAPM beta, a popular measure of systematic risk. SRISK thus does not seem to suffer from some of the criticism Benoit et al. directed towards e.g. the MES measure, which they found unable to separate systemic risk from systematic risk.

I also find that different methods for estimating the long-run marginal expected shortfall (LRMES) component of the SRISK measure have very different dynamics regarding their effect on the rank correlations between SRISK and its components. The correlation dynamics observed with the static LRMES estimator, which is a function of the CAPM beta, introduce procyclicality to the SRISK measure and may also hamper its ability to separate systemic and systematic risk. Furthermore, the absolute magnitudes of the SRISK figures estimated with different estimators of the LRMES may differ as much as tens of billions of U.S. dollars, which is alarmingly high. These considerations should direct the attention of further research to the sensitivity of SRISK to the LRMES model specification and the theoretical and empirical validation of the chosen models.

Based on the analysis of the relationship between SRISK and the current regulatory approach for identifying systemically important financial institutions, as well as on theoretical analysis of the factors affecting SRISK, I can also conclude that SRISK may not be sufficiently able to incorporate the systemic risk effects stemming from the interconnectedness of the financial system. As a possible solution to this problem, I propose further inquiry into network-based measures as addition to the SRISK model specification studied here.

References

Abbe, E. A., Khandani, A. E., and Lo, A. W. (2012). Privacy-preserving methods for sharing financial risk exposures. *American Economic Review*, 102(3):65–70.

Abdi, H. (2007). The Kendall rank correlation coefficient. In Salkind, N., editor, *Encyclopedia of Measurement and Statistics*. SAGE Publications, Thousand Oaks, CA.

Acharya, V. V., Engle, R. F., and Pierret, D. (2014). Testing macroprudential stress tests: The risk of regulatory risk weights. *Journal of Monetary Economics*, 65:36–53.

Acharya, V. V., Engle, R. F., and Richardson, M. (2012). Capital shortfall: A new approach to ranking and regulating systemic risks. *American Economic Review*, 102(3):59–64.

Acharya, V. V., Pedersen, L. H., Philippon, T., and Richardson, M. (2017). Measuring systemic risk. *Review of Financial Studies*, 30(1):2–47.

Adrian, T. and Brunnermeier, M. K. (2016). CoVaR. *American Economic Review*, 106(7):1705–1741.

Alesina, A. and Giavazzi, F. (2010). *Europe and the Euro*. University of Chicago Press, Chicago, IL.

Allen, F. and Gale, D. (2000). Financial contagion. *Journal of Political Economy*, 108(1):1–33.

Basel Committee on Banking Supervision (2013). Global systemically important banks: updated assessment methodology and the higher loss absorbency requirement.

Benoit, S., Colliard, J.-E., Hurlin, C., and Pérignon, C. (2017). Where the risks lie: A survey on systemic risk. *Review of Finance*, 21(1):109–152.

Bhattacharya, S. and Gale, D. (1987). Preference shocks, liquidity, and central bank policy. In Barnett, W. and Singleton, K., editors, *New Approaches To Monetary Economics*, pages 69–88. Cambridge University Press, New York, NY.

Bisias, D., Flood, M., Lo, A. W., and Valavanis, S. (2012). A survey of systemic risk analytics. *Annual Review of Financial Economics*, 4(1):255–296.

Brownlees, C. T., Chabot, B. R., Ghysels, E., and Kurz, C. J. (2015). Back-testing systemic risk measures during historical bank runs. Federal Reserve Bank of Chicago, Working Paper 2015-09.

Brownlees, C. T. and Engle, R. F. (2017). SRISK: A conditional capital short-fall measure of systemic risk. *Review of Financial Studies*, 30(1):48–79.

Chen, C., Iyengar, G., and Moallemi, C. C. (2013). An axiomatic approach to systemic risk. *Management Science*, 59(6):1373–1388.

Curry, T. and Shibut, L. (2000). The cost of the savings and loan crisis: Truth and consequences. *FDIC Banking Review*, 13(2):26–35.

Diamond, D. W. and Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy*, 91(3):401–419.

Engle, R. F. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350.

Engle, R. F. (2009). *Anticipating correlations: A new paradigm for risk management.* Princeton University Press, Princeton, NJ.

Field, A., Miles, J., and Field, Z. (2012). *Discovering Statistics Using R.* SAGE Publications, Thousand Oaks, CA.

Financial Stability Board (2016). 2016 list of global systemically important banks (G-SIBs).

Hansen, L. P. (2012). Challenges in identifying and measuring systemic risk. National Bureau of Economic Research, Working Paper 18505.

Howells, P. and Bain, K. (2008). *The economics of money, banking and finance.* Pearson, London, 4th edition.

Huang, X., Zhou, H., and Zhu, H. (2012). Systemic risk contributions. *Journal of Financial Services Research*, 42(1):55–83.

Idier, J., Lamé, G., and Mésonnier, J.-S. (2014). How useful is the marginal expected shortfall for the measurement of systemic exposure? A practical assessment. *Journal of Banking & Finance*, 47:134–146.

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, 30(1-2):81–93.

Kim, J.-B., Li, L., Ma, M. L. Z., and Song, F. M. (2016). CEO option compensation and systemic risk in the banking industry. *Asia-Pacific Journal of Accounting & Economics*, 23(2):131–160.

Korinek, A. (2011). Systemic risk-taking: Amplification effects, externalities, and regulatory responses. European Central Bank, Working Paper 1345.

Lehar, A. (2005). Measuring systemic risk: A risk management approach. *Journal of Banking & Finance*, 29(10):2577–2603.

Leitner, Y. (2005). Financial networks: Contagion, commitment, and private sector bailouts. *The Journal of Finance*, 60(6):2925–2953.

Lucas, R. E. (1976). Econometric policy evaluation: A critique. In *Carnegie-Rochester conference series on public policy*, volume 1, pages 19–46. Elsevier, New York, NY.

Nucera, F., Schwaab, B., Koopman, S. J., and Lucas, A. (2016). The information in systemic risk rankings. European Central Bank, Working Paper 1875.

Patton, A. J. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review*, 47(2):527–556.

Pearson, K. (1895). Note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, 58:240–242.

Pierret, D. (2015). Systemic risk and the solvency-liquidity nexus of banks. *International Journal of Central Banking*, 11(3):193–227.

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

RStudio Team (2016). *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA.

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1):72–101.

The Volatility Institute (NYU Stern) (2017). The volatility laboratory: Systemic risk analysis. Retrieved from `https://vlab.stern.nyu.edu/` on February 7, 2017.

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York, NY.

Wu, D. and Zhao, X. (2014). Systemic risk and bank failure. Office of the Comptroller of the Currency, Working paper.

Zawadowski, A. (2013). Entangled financial systems. *Review of Financial Studies*, 26(5):1291–1323.